The roles of Polycomb-Group proteins in the development of *Arabidopsis thaliana*

Frazer Thorpe

Supervisor: Dr. W, J. Goodrich

PhD

University of Edinburgh

2008

I. Abstract

The roles of Polycomb-Group proteins in the development of *Arabidopsis thaliana*

The Polycomb-group (Pc-G) complex determines animal cell fate by regulating the expression of the homeotic genes that specify the body pattern. Several Pc-G proteins form a complex, termed Polycomb Repressive Complex 2 (PRC2), which can methylate histone tails. In plants, cell fates are less rigidly determined, suggesting the Pc-G play minor roles in development, or that its effects are more readily reversible. In *Arabidopsis*, there are three homologues of the catalytic unit of the PRC2, encoded by the *MEDEA* (*MEA*), *CURLY LEAF* (*CLF*) and *SWINGER* (*SWN*) genes. *CLF* and *SWN* are expressed throughout development, whereas *MEA* is confined to seed development. The *swn-* mutants appear normal, *clf-* mutants are early flowering, but *swn- clf-* double mutants are only viable in tissue culture, and develop into immortal callus-like material. This suggests *CLF* and *SWN* function is masked by redundancy. *SWN* is widely conserved in flowering plants, suggesting it may have functions independent of *CLF*. The severity of the *swn- clf-* phenotype indicated the Pc-G might play broad roles in plant development, but few targets are known. The aims of this thesis were to determine whether *SWN* had discrete functions in development, and uncover target genes and developmental pathways controlled redundantly by *CLF* and *SWN*. Phenotypic analysis suggested that *SWN* is required to promote the juvenile to adult phase transition and repress leaf initiation rate. Microarray analysis was performed and combined with published "ChIP on chip" data of genome wide loci possessing histone 3 lysine 27 tri-methylation ($H3K27me^3$), a modification specifically catalyzed by the PRC2. My results suggest that there are over 1000 direct targets of Pc-G in *Arabidopsis*, and these tended to be the most highly over-expressed genes in Pc-G mutants. Validation of the microarray data, and phenotypic analysis revealed Pc-G complex is a global regulator of development, and *SWN* and *CLF* play novel roles in stem cell maintenance, promoting and repressing flowering, and in confining embryogenic traits to seed development.

II. Acknowledgements

Hello, I would very much like to my give many thanks to many peoples who have helped and enabled me to achieve the work presented here both socially and scientifically. Firstly, I need to thank my supervisor, Justin, a prince among men, who has shepherded through the joy that was my PhD. The lovely Kim for all the mental support and nourishment, as well as actual nourishment, will not be forgotten, and I thank you muchly! My family should be thanked too, for everything a family should be! And helping me to accomplish this thing, this beast called a PhD, and I thank all them sincerely for the financial side of things and for generally being them. Everyone I have ever shared a beer, or shared too many beers with (Manu, Oli, and Ross (in alphabetical order)); with for the memories of Edinburgh I wanted when I first came here!

Thanks to you all.

Frazer.

III. Abbreviations

A	Adenine
bp	base pair(s)
C	Cytosine
ChIP	chromatin immuno-precipitation
Chip	microarray gene chip
DNA	deoxyribonucleic acid
dCAP	derived cleaved amplified polymorphism
dNTP	any dideoxyribonucleic acid
g	gram
g	gravitiy
GFP	green fluorescent protein
GUS	β-glucuronidase
H	histone
IKI	iodine potassium iodine
K	lysine
kb	kilobase
rpm	revolutions per minute
m	milli (10^{-3})
M	molar
n	nano (10^{-6}
PCR	polymerase chain reaction
PI	propidium iodide
Q-RT-PCR	quantitative reverse transcription polymerase chain reaction
RNA	ribonucleic acid
RNAase	ribonuclease
RT-PCR	reverse transcription polymerase chain reaction
T	thymine
TSA	trichostatin A

u	unit
v	volume
µ	micro (10^{-6})

IV. Contents

1. 0. Introduction

1. 1. Introducing the role of epigenetic control of multi-cellular development

A central problem in multi-cellular development is how whole organisms can develop from a single fertilized egg cell to give rise to many different cell, tissues, and organ types. It is, however, evident that cells of early embryos become different from one another in a process termed patterning. For example in plants, the first cell division of the zygote (fertilized egg cell) divides asymmetrically giving a small apical cell and a basal cell, which are easily distinguishable (Mayer *et al.*, 1993). This very early patterning event is characterized by short range signalling by the phyto-hormone auxin, whose distribution is regulated by polar transport, resulting in a gradient of auxin concentration between cells (Friml *et al.*, 2003). The early patterning events cause cells to acquire localized gene expression patterns, often controlled by "master regulators" such as homeotic genes which are key to specifying differentiation and developmental programs that distinguish the cell, tissue, and organ types of the mature organism. The pattern of gene expression is then required to be "remembered" or to be maintained in on/off states, through cell division as the organism grows. One method to achieve this is through cell intrinsic mechanism, by which genes are maintained in an on/off state inherited through cell division; this can be accomplished through epigenetic mechanisms. The term epigenetic can be used to describe regulation of gene expression through cell division by stable but reversible means but without changing DNA sequence. This is commonly achieved by DNA methylation or chromatin modification. In animals, the Polycomb Group complex (Pc-G) proteins maintain cell fate by locking into place the on/off states of gene activity, from the early stages after patterning and through cell lineages (Simon, 1995; Simon & Tamkun, 2002; Francis & Kingston, 2001). These cell fates are maintained throughout development, i.e. the cell determination of animal cells is rigid and not readily changed. In plants, cell fate determination is comparatively plastic whereby patterned or differentiated tissues can be induced to change fate by becoming de-differentiated and acquiring a different cell fate,

and can form a new plant from already patterned tissues (Ikeda-Iwai *et al.*, 2002; van den Berg *et al.*, 1995). The relatively flexible nature of plant tissue differentiation suggests that epigenetic mechanisms are of less importance in maintaining cell fate compared to animals. However, genetic screens have revealed that plants also possess homologous components of the animal Pc-G, and have developmentally important roles. Therefore, I aim to discover what roles the Pc-G plays in plant development, and what its targets are.

1. 2. Epigenetic changes occur through chromatin modifications

The compaction of eukaryotic DNA is a necessity due to the length of the DNA molecule, for example, the linear DNA molecules in a human cell would be over a metre in length, and has to be packaged into a nucleus only a few micrometers wide. DNA is packaged onto nucleosomes, which are comprised of 145-147 base pairs of DNA coiled around an octomer of core histones with a variable linker region of 20-80 base pairs. This structure gives the appearance of "beads on a string" when viewed under Scanning Electron Microscopy (SEM) at particular salt conditions (Olins & Olins, 1974). The histone octomer is comprised of two copies of H2A, H2B, H3, and H4, and their histone tails are believed to protrude from the nucleosome (Ausio *et al.*, 1989) (Fig. 1. 1). Cytological studies revealed two morphologically distinct forms of chromatin, termed heterochromatin and euchromatin. Heterochromatin stains more intensely with DNA binding stains such as 4'-6-Diamidino-2-phenylindole (DAPI) (Lin & Alfi, 1976), this suggested that it is more densely compacted than the less dense euchromatin. Heterochromatin is composed of highly compacted nucleosomes, giving a dense structure throughout the cell cycle, and is frequently associated with the laminar of the nuclear envelope, telomeres and pericentric regions of chromosomes (Tohno & Tohno, 1993; Elgin & Grewal, 2003; Sun *et al.*, 2001). Heterochromatin frequently contains highly repetitive sequences and is associated with repressed transcription, low gene density, and evenly spaced nucleosome density (as judged by nuclease digestion) (Tohno & Tohno, 1993; Elgin & Grewal, 2003; Sun *et al.*, 2001). Euchromatin is gene rich with irregularly spaced nucleosome packing and is less dense during interphase of

the cell cycle. Euchromatic regions are commonly associated with actively transcribed regions of DNA (Goodrich & Tweedie, 2002; Hsieh & Fischer, 2005; Martienssen & Colot, 2001).

The histone tails are highly conserved between eukaryotic kingdoms, from plants to humans, although they appear not to play a central role in nucleosome structure, as structural integrity is maintained when histone tails are removed. Their conservation was considered a mystery (Cheah & Osborne, 1977; Olins & Olins, 1974; Jenuwein & Allis, 2001). The paradox of the broad conservation of histone tails, yet them being of no direct structural benefit to chromatin structure was resolved with the discovery that histone tails can be covalently modified. This feature of histone tails can play a significant role in regulating gene expression via an epigenetic mechanism, which is often correlated with different chromatin structure. There are two modifications to chromatin that are deemed epigenetic, (1) DNA methylation, which in plants is very stable, can be inherited between generations, and is largely involved in genome defence (Waterhouse *et al.*, 2001; Matzke & Matzke, 1998), but shall not be further discussed. (2) Histone tail modification, for example, the animal Pc-G complex can tri-methylate lysine 27 on histone 3 ($H3K27me^3$) and this is known to correlate with a repressed gene state (Kahn *et al.*, 2006). Whereas other histone marks are found commonly associated with active transcription, such as $H3K4me^3$ (Jenuwein & Allis, 2001). In addition to this, the chromatin type where the epigenetic marks are found can determine the transcription activity of a gene, for example the $H3K27me^3$ mark is correlated with gene repression in euchromatin, but can be associated with transcription reactivation of transposable elements and repeat elements in heterochromatin (Tariq *et al.*, 2003). Many histone modifications are possible and include methylation, acetylation, phosphorylation, ubiquitination, and sumoylation (Jenuwein & Allis, 2001). Histone modification status of a gene can also vary depending on cell cycle stage environmental influence suggesting a dynamic process but how they are targeted is presently unknown. Transcription activity is affected not only by the presence of these marks, but also their location and association with other types of marks. For example heterochromatin in

plants is associated with H3K27me^2 and H3K9me^2 together, and DNA methylation (Jackson *et al.*, 2002). These observations led to the histone code hypothesis that suggests that histone modifications convey information that specifies gene expression activity, and other features such as DNA repair (Jenuwein & Allis, 2001). The interpretation of each of the modifications depends on the presence of, and combination with, many other histone modifications (Jenuwein & Allis, 2001). This phenomenon can be used as a cellular memory, to lock in place a transcriptional profile of clonally distinct cell lineages, tissue types, or developmental phases (Elgin & Grewal, 2003; Sung *et al.*, 2006). Of particular interest is the methylation of H3K27, which has been shown to be mono-, di-, or tri-methylated (Jenuwein & Allis, 2001).

1. 3. The Pc-G genes mediate cell fate determination

1. 3. 1. The animal Pc-G genes have a common role in developmental memory

The cellular organization of both plants and animal embryos is integral to the future viability, health, and vitality of the adult. This is controlled via a number of interconnected pathways involving both genetic and epigenetic methods to accurately choreograph patterning and differentiation. One epigenetic pathway common to both plants and animals is the Pc-G, which regulates gene expression in response to both internal and external signalling. The *Drosophila* Pc-G complex was the first to be identified and is well characterised. This topic has been extensively reviewed (Beuchle *et al.*, 2001; Breiling & Orlando, 2002; Dejardin & Cavalli, 2005; Francis & Kingston, 2001) and only a brief overview will be outlined here.

The basic body plan of the fly is determined during the early stages of *Drosophila* development by the spatial expression of homeotic genes. The embryo develops into the larval stage where it is divided along the anterior-posterior axis into repeated units termed imaginal discs or para-segments. These are distinguished from one another by different combinations of homeotic gene expression and activity which are maintained throughout cell division cycles, and in specific cell lineages, throughout development

(Elgin & Grewal, 2003; Sung *et al.*, 2006). Homeotic gene expression is spatially restricted to domains along the anterior-posterior axis to determine the identity of discrete fly body segments in adult flies (Scott & O'Farrell, 1986). For example, the homeotic gene *ABDOMINAL-A* (*abdA*) is expressed discretely in the abdomen, and not in the head or thorax domains (Shimell *et al.*, 1994). The expression of homeotic genes is initiated after the first two hours of embryogenesis and is maintained until death (Scott & O'Farrell, 1986). Homeotic gene pattering is initiated by the transient expression of DNA binding transcription factors encoded by gap and pair-spacing genes, whose expression is restricted to specific domains of the embryo along the anterior-posterior axis (Scott & O'Farrell, 1986). *abdA* expression is repressed in the anterior segments due to the presence of the gap genes *KRUPPEL* and *HUNCHBACK* (*HBK)* (Shimell *et al.*, 1994; White & Lehmann, 1986; Harding & Levine, 1988; Reinitz & Levine, 1990; Qian *et al.*, 1993). The homeotic gene *ULTRABITHORAX* (*UBX*) is also repressed by *HBK* in the anterior domains (Zhang & Bienz, 1992; Qian *et al.*, 1993). The expression of the gap and pair-rule genes is lost four hours into embryogenesis, yet the expression patterns of homeotic genes such as *UBX* is maintained in the posterior domains and is stably repressed in anterior domains (Fig. 1. 2). This raises the question of how these patterns of activity are maintained through cell division in the absence of these regulators that determine on/off states of gene expression.

The Polycomb-Group (Pc-G) and Trithorax-Group (Trx-G) are a large group of genes identified genetically on the basis of anterior-posterior defects. Small changes in Pc-G activity, for example a heterozygous *POLYCOMB* (*Pc*) mutant (*Pc-/+*), gives a characteristic homeotic conversion of an extra sex-comb phenotype in adult male flies. Wild-type male flies have a tuft of hair on the prothoracic (most anterior) legs, mutants have sex-combs on the mesothoracic and metothoracic (more posterior) legs (Hannah-Alava, 1958). This is due to the ectopic expression of the homeotic gene *SEX COMB* in the meso- and metothoracic leg imaginal discs (Glicksman & Brower, 1988). Further reduction of Pc-G activity results in early embryo lethality, and embryos show severe mis-expression of homeotic genes. For example, homozygous mutants of the

ENHANCER OF ZESTE (*E(z)*) initially show a wild-type pattern of homeotic genes expression, determined by the gap and pair-spacing genes but after their decay homeotic genes become de-repressed after four hours. An example of which is *UBX*, which is normally confined to the posterior regions but is found ectopically expressed in the anterior segments of *e(z)*- mutants, resulting in their mis-specification (Kahn *et al.*, 2006) (Fig. 1. 2). This indicates that the Pc-G genes are required to maintain the repressed state of homeotic genes outside of regions where they are normally expressed.

The Trx-G maintains the expression of homeotic genes in the active expression state as demonstrated when members of the Trx complex are mutated. For example, *trx*- mutants show reduced expression of *UBX* resulting in the mis-specification of the posterior segments (Farkas *et al.*, 1994). The Pc-G and Trx-G complexes are expressed ubiquitously throughout the embryo to maintain the prescribed homeotic gene expression in the different segments (Ringrose & Paro, 2004). The dramatic and embryo lethal phenotypes observed in these mutants illustrate the importance of the Trx-G and Pc-G in regulation of homeotic transcription factors.

1. 3. 2. The animal Pc-G genes act in complexes

The molecular cloning of the Pc-G members shows they encode structurally unrelated proteins; the explanation of their similar functions was uncovered by biochemical studies, showing that they act together in two distinct complexes. These complexes are termed the Polycomb Repressive Complex1 (PRC1) and PRC2. The four core components of the *Drosophila* PRC2 are ENHANCER OF ZEST (E(z)), SUPPRESSOR OF VARIATION12 (Su(z)12), EXTRA SEX COMB (Esc), and P55. These proteins have been shown to physically and genetically interact to repress homeotic gene expression in the *Drosophila* embryo (Fig. 1. 3) (Beuchle *et al.*, 2001). The E(z) proteins contains a SET domain. The SET domain was found to convey histone methyltransferase catalytic ability, in a variety of evolutionarily conserved SET domain proteins related to *Su(var)39* in *Drosophila* (Rea *et al.*, 2000). As the catalytic unit of the PRC2 *E(z)* maintains gene repression by the adding di- and tri-methyl groups to lysine

27 on histone 3 ($H3K27me^2$ and $H3K27me^3$) at target gene loci. This was plainly demonstrated via *in vitro* assays and by the loss of the $H3K27me^3$ mark at target gene loci *in vivo* in *e(z)* mutants (Czermin *et al.*, 2002; Cao & Zhang, 2004). The SET domain and flanking cysteine rich regions are thought to create a groove in which the histone tail lays on one face and a S-adenosyl-L-methionine cofactor binds to the opposite face. The lysine residue approaches the co-factor in the centre of the groove, and the transfer of methyl groups occurs (Qian & Zhou, 2006).

The other components of the PRC2 complex are essential for *E(z)* function, as *E(z)* is not capable of binding DNA or histone tails without the other Pc-G members (Czermin *et al.*, 2002; Cao & Zhang, 2004). For example, Su(z)12 encodes a VEFS domain containing protein that interacts directly with E(z) and is essential for its function in patterning the *Drosophila* embryo (Yamamoto *et al.*, 2004). Esc also physically interacts with E(z) and encodes a WD-repeat protein (Ng *et al.*, 2000). P55 physically interacts with Esc in the Pc-G and also encodes a WD-repeat domain protein with histone deacetylase catalytic properties (Tie *et al.*, 2001). The interactions of these four components form the Pc-G complex (Fig. 1. 3). In addition to the Pc-G, P55 is also a member of the Chromatin Assembly Factor 1 complex (CAF1) (Bowen *et al.*, 2004; Tie *et al.*, 2001). This potentially links histone modification and chromatin assembly in co-ordinating correct gene regulation.

The second Pc-G complex, PRC1, interprets the marks made by the PRC2, particularly $H3K27me^3$, to bring about stable repression of target genes. The PRC1 complex is comprised of PLEIOHOMEOTIC (PHO), POLYCOMB (PC), POSTERIOR SEX COMBS, SEX COMBS ON MIDLEG, and POLYCOMB-LIKE (Schwartz & Pirrotta, 2007). Pho binds to target gene DNA and recruits Pc, which co-localises with the other members of the PRC1 (Mohd-Sarip *et al.*, 2006). The binding of the PRC1 to target loci maintains gene repression, but how it does this is still unclear. It has been suggested that remodelling target chromatin by PRC1 makes DNA inaccessible to the transcriptional machinery ensuring stable repression. This is a possible proposition as Su(z)12 interacts with HETEROCHROMATIN PROTEIN1 (HP1) which is needed to stably repress

known Pc-G targets (de la Cruz *et al.*, 2007). In addition, HP1 is found to interact with other repressive marks such as H3K9me^3 to maintain a repressed expression state (Jacobs & Khorasanizadeh, 2002; Lachner *et al.*, 2001).

The Pc-G and Trx-G complexes bind to their DNA targets through Polycomb Response Elements (PRE) and Trithorax Response Elements (TRE), respectively, and these may or may not be mutually exclusive (Ringrose & Paro, 2007). It is believed that the Pc-G complexes are targeted to the PREs by sequence specific DNA binding transcription factors, as the components of the Pc-G themselves are unable to bind to DNA. Mutating **a PRE causes that gene loci to become mis-regulated, for example mutating a PRE at the *UBX* loci results in *UBX* being found ectopically expressed in the anterior regions of the embryo and results in mis-specification of these segments into posterior segments** (Sipos *et al.*, 2007). Genome-wide analysis, using Chromatin Immuno-Precipitation, DNA microarray gene chip (ChIP on chip) techniques and polytene chromosomes, has revealed *Drosophila* possesses 100-200 PREs (Ringrose *et al.*, 2003; Tolhuis *et al.*, 2006; Zink *et al.*, 1991). In addition to homeotic gene regulation the Pc-G and Trx-G complexes regulate a range of other process and genes such as cell cycle regulation, and cell signalling (Schuettengruber *et al.*, 2007).

Stable repression of genes by epigenetic means in animals is thought to be important for maintenance of a defined differentiation state. It is therefore necessary that gene repression be maintained throughout cell divisions. This can be achieved in two ways, either, that targets are repressed by heterochromatin, or, changes in chromatin structure leads to gene repression, which is maintained through cell division and maintained in the following generation of cells (Costa & Shaw, 2006; Exner & Hennig, 2008). The exact nature of these processes to regulate gene expression is unclear and the mechanisms by which nucleosome modifications are inherited remain a topic of debate.

1. 4. The Pc-G regulates *Arabidopsis* development

Arabidopsis thaliana has been employed as a model species for plant genetics due to a number of adventitious features, including being a small, self-pollinating plant with a rapid life cycle whose small genome has been fully sequenced. These attributes make it extremely useful for investigation of the genetic processes that regulate development, and in the future this knowledge may be transferred to agronomically important species.

The life cycle of *Arabidopsis* is categorized into five stages, embryonic, juvenile, adult, reproductive, and senescence (Fig. 1. 4). Each of the phases are characterized by distinct changes in growth forms and gene expression. Embryogenesis, the transitions between Juvenile-Adult (J-A), Adult-Reproductive (A-R), and the regulation of meristematic regions are controlled genetically and epigenetically by the plant Pc-G complex.

In a similar way the Pc-G components were first discovered in *Drosophila*, the plant Pc-G members were identified in genetic screens for developmental regulators. Strikingly, the molecular isolation of plant Pc-G components revealed that nearly all encoded homology of the PRC2. The homologues of the Pc-G in *Arabidopsis* play similar roles in development to those in *Drosophila.* As described above, there are four central components of the PRC2 complex in flies: *E(z), Su(z)12, Esc*, and *P55*, and each of the genes have homology to genes in *Arabidopsis*. For example, *Esc* is homologous to *FERTILIZATION INDEPENDENT ENDOPERM* (*FIE*) (Guitton *et al.*, 2004; Luo *et al.*, 1999; Ohad *et al.*, 1999). While *P55* is homologous to *MULTI SUPPRESSOR OF IRA1* (*MSI1*) (Luo *et al.*, 1999) (Kohler *et al.*, 2003a). The other PRC2 components are encoded by small gene families; *FERTILIZTION INDEPENENT SEED2* (*FIS2*), *VERNALIZATION2* (*VRN2*), and *EMBRYONIC FLOWER2* (*EMF2*) are homologues of *Su(z)12* (Luo *et al.*, 1999; Gendall *et al.*, 2001; Chanvivattana *et al.*, 2004; Birve *et al.*, 2001). *MEDEA* (*MEA*), *CURLY LEAF* (*CLF*), and *SWINGER* (*SWN*) are homologues of *E(z)* (Goodrich *et al.*, 1997; Chanvivattana *et al.*, 2004; Grossniklaus *et al.*, 1998) (Fig. 1. 3). Interestingly, there are no proteins in the *Arabidopsis* genome homologous to the PRC1 proteins including the complexes' namesake *Pc* (Chanvivattana *et al.*, 2004). The

duplication of components and the lack of PRC1 complex indicates that although there are similarities between Pc-G regulated gene expression in *Drosophila* and *Arabidopsis*, the mechanisms by which they achieve gene repression must differ.

Genetic and molecular evidence suggests that the duplication of several Pc-G members has allowed the evolution of several plant PRC2-like complexes with partially discrete roles in development. The nature of these complexes and their roles are briefly discussed below.

1. 4. 1. The FIS2-Pc-G complex represses endosperm proliferation in seed development

In *Arabidopsis*, as in other flowering plants, zygotic development occurs as a result of the characteristic double fertilization event. The seed develops from a fertilized ovule, and contains the two zygotic tissues within the surrounding maternal seed coat tissues. Following the double fertilization event in *Arabidopsis* flowers, the fertilized egg cell in the ovule divides asymmetrically and the apical cell continues to divide. First into two cells, then four, and eight cells. The embryo continues to divide and forms the globular stage and heart stage embryo. At the heart stage all the embryonic tissue types have been determined, the cotyledons, the shoot apical meristem (SAM), root apical meristem (RAM), and the vasculature. The following embryo stages are a morphological extension of this stage and termed the torpedo stage and the walking stick stage (Fig. 1. 4). While the egg cell of the gametophyte is forming the embryo, the central cell of the female gameteophyte also undergoes divisions. The central cell divides into an ephemeral network of cells termed the endosperm, which provides a nutrient source to the developing embryo. The endosperm is degraded in the later stages of embryogenesis, the embryo then fills the developing seed, and the cotyledons then carry out the storage capacity of the dissipated endosperm (Fig. 1. 4) (Park & Harada, 2008; Willemsen & Scheres, 2004; Mayer *et al.*, 1993).

The fertilization independent seed (FIS) genes share similar mutant phenotypes and were found in genetic screens for seed development in the absence of fertilization. Whereas, wild-type ovules abort if unfertilized, in the FIS class mutants the endosperm central cell proliferates to form endosperm tissue. Emasculated FIS class mutants, such as *mea-/+*, *fis2-/+*, *msi1-/+*, and *fie-/+* plants, 50% of the central cells are mutant and show autonomous seed set and produce an over-proliferation of deformed opaque endosperm that eventually aborts, and no embryo is present (Grossniklaus *et al.*, 1998; Chaudhury *et al.*, 1997; Guitton *et al.*, 2004; Kiyosue *et al.*, 1999; Luo *et al.*, 1999; Ohad *et al.*, 1996).

MSI1, *FIE*, *FIS2*, and *MEA* are expressed in the central cell of the female gametophyte to regulate seed set, and endosperm proliferation but *FIS2* and *MEA* are not found in the embryo (Wang *et al.*, 2006; Ohad *et al.*, 1999; Luo *et al.*, 2000). However, a second characteristic of the FIS mutations i.e. *MEA*, *FIE*, *FIS2*, or *MSI1* genes is that when the mutant ovule is fertilized it results in embryo lethality, in a maternal specific manner. For example if a seed receives a mutant maternal copy of *mea-* the seed aborts irrespective of the paternal copy of *MEA* (Grossniklaus *et al.*, 1998; Nowack *et al.*, 2007). Mutants that possess a maternally inherited copy of *mea-* in central cell gametophytes, when fertilized, causes embryos to show disorganized patterning of root, shoot and vasculature tissue in the early embryo compared to wild-type embryos (Grossniklaus *et al.*, 1998). At the heart stage of embryogenesis the seeds are aborted in an unknown mechanism, but probably due to defected endosperm. The reasoning for the abortion is believed to be a consequence of "parent of origin conflict", whereby *MEA* is required to suppress the contribution of paternal expression (Grossniklaus *et al.*, 1998). This demonstrates that *MEA*, and the other FIS class mutants, are required to repress endosperm proliferation prior to fertilization, and to allow proper embryo development through correct endosperm development.

The *mea-* phenotype is partially attributed to ectopic expression of the MADs box transcription factor gene *PHERES1 (PHE1)*, one of its probable many targets, which was identified in a microarray study (Kohler *et al.*, 2003b). The *mea- phe1-* double

mutant partially rescues the *mea-* mutant (Kohler *et al.*, 2003b). The phenotype of *mea-* is identical to *fie-*, *fis2-*, and *msi1-* which supports the hypothesis that these genes act in the same complex (Chaudhury *et al.*, 1997; Kohler *et al.*, 2003a). In addition, MEA-FIS2, FIS2-FIE and FIE-MSI1 were found to physically interact (Fig. 1. 3) (Katz *et al.*, 2004; Yadegari *et al.*, 2000; Luo *et al.*, 2000; Spillane *et al.*, 2000).

SWN is also expressed in the central cell of the ovule before fertilization and then in the embryo proper, but *swn-* mutants do not show any seed defects (Wang *et al.*, 2006). However, the *swn- mea-* double mutant has a mildly enhanced *mea-* phenotype, as *swn- mea-* plants show increased autonomous seed set when unfertilized (Wang *et al.*, 2006). SWN was found to interact with FIS2 and FIE in Yeast-two-Hybrid (Y-2-H) assay (Wang, *et al*., 2006; Chanvivattana, *et al.*, 2004. This indicates that *SWN* and *MEA* act partially redundantly in the FIS2-Pc-G to repress central cell proliferation prior to fertilization.

Differences in expression patterns of SWN and MEA/FIS2 suggest that only certain aspects of embryogenesis are regulated redundantly by *SWN* and *MEA* in the endosperm. The *clf- mea-* mutant embryos and seeds show a *mea-* phenotype (personal communication Dr. U. Grossniklaus) and *clf-* embryos and seeds show a wild-type morphology. This illustrates that *MEA* and *SWN* are capable of acting partially redundantly but *CLF* and *MEA* may not.

1. 4. 2. The EMF2-Pc-G complex represses flowering

MEA and *FIS2* expression is confined to embryogenesis. This gives rise to the hypothesis that the *MEA-FIS2-FIE-MSI1*-Pc-G complex acts specifically during embryogenesis. The *SWN, CLF, VRN2, EMF2, FIE,* and *MSI1* genes are expressed during embryogenesis and throughout development, suggesting these act as a post-germination Pc-G complex (Fig. 1. 3). Following maturation and germination of the seed, a distinctive transcriptional profile is in action to regulate development after

germination. The post-germination development is also regulated by the Pc-G, particularly flowering time, vernalization, and organogenesis.

Flowering is one of the most important "decisions" a plant has to make, and as such, it is under exceptionally rigorous genetic control. *Arabidopsis* possesses a mighty arsenal for responding to external and internal cues that induce or repress flowering. There are four known flowering pathways in *Arabidopsis;* these include the vernalization response, light dependent pathway, autonomous pathway, and the gibberellic acid (GA) pathway (Fig. 1. 5) (Putterill *et al.*, 2004; Simpson, 2004; Simpson *et al.*, 1999; Parcy, 2005; Sablowski, 2007a). Of particular interest for this thesis are the autonomous and vernalization pathways (Fig. 1. 5.).

Given the correct environmental (light) and endogenous (such as GA) signals it causes the transcriptional activation of the floral integrators *SUPPRESSOR OF CONSTANTS* (*SOC1)* a MADS-box transcription factor), *FLOWERING LOCUS T* (*FT*), and *LEAFY* (*LFY*) expression (Fig. 1. 5.) (Hayama & Coupland, 2003; Imaizumi & Kay, 2006; Eriksson *et al.*, 2006). Over-expressing *LFY* and/or *FT* results in early flowering (Kardailsky *et al.*, 1999; Blazquez *et al.*, 1997; Weigel & Nilsson, 1995), and mutating *LFY*, *FT* or *SOC1* causes late flowering (Weigel *et al.*, 1992; Kobayashi *et al.*, 1999; Borner *et al.*, 2000; Lee *et al.*, 2000) (Fig. 1. 5). *FT* and *LFY* promote *APETALA1* (*AP1*), and *CAULIFLOWER* (*CAL)* expression in the SAM, triggering the switch from the vegetative meristem to the inflorescence meristem (IM) (Liljegren *et al.*, 1999; William *et al.*, 2004; Parcy *et al.*, 1998; Ruiz-Garcia *et al.*, 1997; Weigel *et al.*, 1992; Wagner *et al.*, 1999). In addition to its role in flowering time *LFY* is required to maintain the IM and floral meristem as *lfy*- mutants show vegetative characteristics (Parcy *et al.*, 1998; Lohmann *et al.*, 2001). Expression of *LFY* and *WUSCHEL* (*WUS*) in the IM induces floral meristem formation through activation of floral organ identity genes like *AGAMOUS* (*AG*) (Lohmann *et al.*, 2001) (Fig. 1. 5). Transcription factors *PISTILATA* (*PI*), *AP2, AP1, AG,* and a range of other genes that specify the different tissue types of the flower as per the ABC model (Coen & Meyerowitz, 1991) are modulated by *LFY,* and other regulatory genes expressed in the floral tissues to define their differentiation

states. For more detailed descriptions of flowering time and floral regulation see reviews (Sung & Amasino, 2004; Robles & Pelaz, 2005; Zik & Irish, 2003; Amasino, 2004; Boss *et al.*, 2004; Simpson *et al.*, 1999; Bastow & Dean, 2003).

The role of the Pc-G in regulating flowering time can be observed in *clf-* mutant plants, which exhibit a strong vegetative phenotype; showing curled leaves, early flowering and partial homeotic transformations of floral organs (Fig. 1. 6) (Goodrich *et al.*, 1997). The *clf-* phenotype is largely a consequence of ectopic expression of *AG* in the leaves and flowers (Goodrich *et al.*, 1997; Chanvivattana *et al.*, 2004). AG is a MADS box transcription factor that is literally central to the ABC floral meristem model. Expression of *AG* is required for the differentiation of central whorls of stamen and carpels. *ag-* mutants show homeotic transformation of these tissues into repeated whorls of sepals and petals (Gomez-Mena *et al.*, 2005; Bowman *et al.*, 1989). Ectopic expression of *AG*, in transgenic plants containing the *35s::AG* construct, induces early flowering, leaf curling, and homeotic transformations of the petals and sepals into stamen and carpeloid structures (Mizukami & Ma, 1992). The *clf- ag-* double mutant shows the homeotic transformations of *ag-* mutants, but the leaf curling and the early flowering time defects in *clf-* plants are almost entirely lost (Goodrich *et al.*, 1997). *EMF2* and *CLF* are likely to regulate relatively common aspects of development as *emf2-* mutants show a similar, but more severe phenotype than *clf-* mutants. The *emf2-* mutants flower almost immediately after germination, without forming juvenile leaves and few if any adult leaves (Chen *et al.*, 1997; Moon *et al.*, 2003b; Chanvivattana *et al.*, 2004). The lack of juvenile leaves of *emf2-* mutants may be due to the de-repression of a potent floral promoter, but it may indicate a role for the Pc-G in maintaining the juvenile phase of development. The flowers of *emf2-* are severely deformed and sterile (Chen *et al.*, 1997; Moon *et al.*, 2003b; Chanvivattana *et al.*, 2004). Expression analysis and the use of floral marker lines showed that, like *clf-*, floral organ identity genes such as *AG* are ectopically expressed are expressed in *emf2-* (Moon *et al.*, 2003b; Chanvivattana *et al.*, 2004).

Co-suppression of *FIE* causes early flowering, floral defects, and ectopic expression of floral organ identity genes, such as *AG*, post-germination (Katz *et al.*, 2004). Co-suppression of *MSI1* causes ectopic expression of *AG* (Bouveret *et al.*, 2006). The combined phenotypic evidence indicates that the *CLF-EMF2-FIE-MSI1*-Pc-G complex plays a central role in repressing flowering time and floral organ identity. This is supported by *in vivo* and Y-2-H data showing that the components of this complex physically interact (Fig. 1. 3) (Chanvivattana *et al.*, 2004; Wood *et al.*, 2006). This strongly suggests this Pc-G complex acts directly to regulate expression of homeotic and flowering time genes.

However, co-suppressed *MSI1* plants show delayed flowering and a range of floral defects (Bouveret *et al.*, 2006). *MSI1* regulates flowering through the regulation of *SOC1* and probably other floral promoters, as *SOC1* expression is reduced in co-suppressed *MSI1* plants. As with *clf*- mutants, *AG* is also found ectopically expressed (Bouveret *et al.*, 2006). *MSI* and *CLF* are likely to regulate flowering time through different pathways as co-suppressed *MSI1 clf*- plants show an intermediate flowering time phenotype compared to the parent mutants.

The hypothesis that the Pc-G is fundamental in regulating flowering time is enhanced by evidence that a specific Pc-G complex co-ordinates the vernalization response.

1. 4. 3. The Pc-G in plants provides an epigenetic memory of winter

The vernalization response is probably one of the best-characterized functions of a repressive Pc-G complex in plants. Vernalization is the response to prolonged cold to permit flowering. Most *Arabidopsis* accessions require vernalization, as they possess functional *FRI* and *FLC* genes. *FRI* promotes the potent floral repressor *FLC* (Fig. 1. 5). The exposure to cold causes decreased *FLC* expression, which permits flowering to occur (Sheldon *et al.*, 1999). The repression of *FLC* in the cold is mediated by Pc-G as it was demonstrated that *VRN2* is the key component of the vernalization response. *vrn2*-plants are vernalization insensitive i.e. they fail to flower after vernalization treatment,

which was found to be a consequence of high *FLC* expression (Gendall *et al.*, 2001). Similar to wild-type plants, *vrn2*- mutants show high *FLC* expression after 1-2 weeks of vernalization, wild-type plants subsequently show reduced *FLC* expression in the following weeks, but *vrn2*- mutants show a transient decreased *FLC* expression but regains its high expression level (Gendall *et al.*, 2001). The resulting increase in *FLC* expression causes delayed flowering or *vrn2*-, explaining the insensitivity to vernalization treatment. This illustrates that the vernalization response is a classic epigenetic phenomenon as the expression state is not inherited through generations but is acquired through environmental cues and it maintained through cell cycles and is very stable. *VRN2* is required for the stable maintenance of *FLC* repression or memory of winter function in plants. VRN2 has been shown to physically interact with both SWN and CLF proteins in Y-2-H assays (Chanvivattana *et al.*, 2004). This strongly indicates that the Pc-G complex is required to mediate the vernalization response.

1. 5. Redundancy masks the role of Pc-G

The *E(z)* homologues *MEA, CLF,* and *SWN* are the focus of this thesis. Based on protein comparisons and DNA sequences *SWN* is believed to be the ancestral *E(z)* gene in plant, whose duplication gave rise to both *MEA* and *CLF* (Spillane *et al.*, 2007). Each of these genes possess the SET domain that is required to catalyze histone methyltransferase activity in *Drosophila* (Czermin *et al.*, 2002), the activity of which appears to be maintained in *Arabidopsis* (Schubert *et al.*, 2006; Lindroth *et al.*, 2004). The functions of these genes has been partially compartmentalized through the different expression patterns, for example *MEA* expression is restricted to seed development (Grossniklaus *et al.*, 1998; Wang *et al.*, 2006). *SWN* is expressed in the central cell prior to fertilization (Wang *et al.*, 2006), and in the embryo throughout embryogenesis, and is found throughout vegetative and reproductive development predominantly in the meristematic tissues (Chanvivattana *et al.*, 2004). This expression pattern of *SWN* is mirrored by *CLF* (Goodrich *et al.*, 1997).

Although *swn-* plants have no documented seed defects, *swn-* enhances aspects of the *mea-* phenotype (Wang *et al.,* 2006), suggesting *SWN* acts with partial redundancy with *MEA,* and *SWN* is of developmental importance. However, *swn-* mutants have no gross developmental defects post-germination (Fig. 1. 6) (Chanvivattana *et al.,* 2004). Considering the protein similarity and expression patterns of *SWN* and *CLF* the double mutant was generated. *swn- clf-* double mutant phenotype massively enhances the *clf-* phenotype (Chanvivattana *et al.,* 2004). Seeds of the *swn- clf-* double mutant are only viable on sterile tissue culture, but following germination, show root defects including a "pickle" root phenotype; being swollen, opaque and stunted in growth, which eventually forms into a callus-like material (Fig. 1. 6). The aerial meristem fails to produce any true leaves and also develops into callus-like tissue (Chanvivattana *et al.,* 2004). The *swn- clf-* phenotype mutation suggests that SWN and CLF act redundantly to potentially regulate many aspects of organogenesis (Chanvivattana *et al.,* 2004). The severity of the double mutant phenotype compared to *clf-* or *swn-* alone is likely a consequence of redundancy because the expression patterns and proteins are almost identical. This has made characterization of the true function of *SWN* and *CLF,* or indeed the Pc-G in post-germinative development extremely difficult. This does illustrate that the entirety of Pc-G function was previously masked by the functional redundancy of Pc-G components, and is seen in the *VRN2* and *EMF2* Pc-G members (Schubert *et al.,* 2005).

This callus-like phenotype is replicated in the *vrn2-1 emf2-3* plants (Schubert *et al.,* 2005). This indicates that there is functional redundancy between both *VRN2* and *EMF2,* and *SWN* and *CLF,* and implies that the Pc-G regulate global organogenesis and morphogenesis through a series of Pc-G complexes that are interchangeable due to the ability of homologous components acting redundantly. The regulation of global organ patterning appears to be carried out by either of the two sets of homologues, as long as either *E(z)* homologue and a *Su(z)12* homologue is present then organogenesis is possible. This remains hypothetical, as no conclusive evidence has yet been generated to show that different combinations of the Pc-G complexes play other roles in

development, exclusively to regulate embryogenesis, flowering time, floral organ identity, and vernalization.

All members of the *Arabidopsis* PRC2 Pc-G complex are expressed in seeds but only *MEA, FIS2, FIE,* and *MSI1* are expressed in the endosperm post-fertilization and show embryonic or seed phenotypes when mutated. Interestingly the other members (*SWN, CLF, VRN2,* and *EMF2)* are expressed in the female gametophyte up until seed maturation, yet do not show seed defects. The role of these members in seed development is unclear. The possibility remains that these members are acting redundantly; alternatively, they may effect gene expression without any visible morphological defects.

Although there appears to be functional redundancy between *SWN* and *CLF,* it has been proven that there is only partial redundancy, as *CLF* acts to repress *AG* expression in the leaves and the outer whorls of flowers independently of *SWN,* hence the difference of *clf-* and *swn-* phenotypes. The independent/discrete role of *CLF,* from *SWN* or *MEA,* was demonstrated by over-expressing *CLF* in a *clf-* background could complement the *clf-* phenotype to a wild-type phenotype, whereas over-expressing *SWN* or *MEA* in a *clf-* background did not effect the *clf-* phenotype (Chanvivattana *et al.,* 2004).

1. 6. The plant Pc-G acts via H3K27me^3 in *Arabidopsis*

SWN and CLF are thought to regulate target genes through the addition and maintenance of H3K27me^3 at target loci, similar to the *E(z)* protein in Drosophila (Czermin *et al.,* 2002; Schubert *et al.,* 2006; Lindroth *et al.,* 2004). Although this is yet to be directly tested *in vitro, in vivo* studies of *swn- clf-* double mutants and *fie*-mutants shows massive loss of H3K27me^3 in euchromatin, but *swn- clf-* and *fie-* mutants show and increased levels of H3K27me^3 in heterochromatin regions (Lindroth *et al.,* 2004). This indicates that a SWN-CLF-Pc-G complex acts to repress gene expression through the H3K27me^3 histone modification.

CLF represses *AG* by the addition of the H3K27me^3 mark to defined locations of the *AG* locus which then spreads in a non-sequence specific manner (Schubert *et al.*, 2006). This is further supported by *clf*- mutants showing reduced H3K27me^3 at the *AG* locus and CLF is found enriched at the *AG* locus (Schubert *et al.*, 2006).

The vernalization response, i.e. the reduced expression of *FLC* after prolonged cold treatment to permit flowering is also regulated by histone methylation. In wild-type plants after 4-6 weeks of cold treatment the *FLC* locus showed increased H3K27me^3 which is correlated to its decreased expression (Gendall *et al.*, 2001; Sheldon *et al.*, 1999). *vrn2*- mutants initially show increased H3K27me^3 mark on *FLC* after ~2 weeks but then H3K27me^3 mark is found reduced over the following weeks coimpared to *VRN2*+ plants. The decrease in methylation correlates with increased expression of *FLC* (Gendall *et al.*, 2001). Therefore, *VRN2* is specifically required to maintain the H3K27me^3 mark on *FLC* and this is likely created by *SWN* and *CLF*. The fact that CLF has been found enriched at the *FLC* loci (personal communication Dr. D. Schubert) provides compelling evidence that the Pc-G complex functions in the vernalization response. However, *clf- FRI*+ plants are still sensitive to vernalization treatment, suggesting the Pc-G mediated H3K27me^3 mark on *FLC* is either a discrete function of *SWN* or the redundant function of *SWN* and *CLF*.

Genome analysis of the location of the H3K27me^3 mark carried out on 10 day old Wild type seedlings discovered that ~4400 distinct single loci possessed H3K27me^3, rather than generalized spreading of the H3K27me^3 mark (Zhang *et al.*, 2007; Turck *et al.*, 2007). It was also noted that islands of H3K27me^3 was principally located in euchromatin, in fitting with the proposed general role to temporally and spatially repress gene expression and the covering of large genomic regions were confined to heterochromatin (Zhang *et al.*, 2007). H3K27me^3 was largely independent of other epigenetic regulation processes such as DNA methylation, and *siRNA* pathways (Chanvivattana *et al.*, 2004). The several distinct differences to *Drosophila*, the lack of spreading of H3K27me^3 and the reduced presence of H3K27me^3 in heterochromatin,

implies there are fundamental differences in the mechanisms by which H3K27me^3 is used and interpreted in plants and animals.

The Pc-G, particularly *SWN* and *CLF*, is fundamental in creating/maintenance of the H3K27me^3 mark as in *swn- clf-* material there is a massive reduction of H3K27me^3 in euchromatic DNA, but a relative increase in H3K27me^3 was found in heterochromatic regions of DNA in the nucleus (Lindroth *et al.*, 2004). This illustrates that *SWN* and *CLF* are required for this mark to be present in euchromatin. Interestingly not all H3K27me^3 was lost. This may be due to incomplete active or passive loss of H3K27me^3, or via the action of other SET domain proteins creating the H3K27me^3 mark. The latter point is supported by there being approximately 32 SET domain proteins in the *Arabidopsis* genome (Baumbusch *et al.*, 2001). They may be important for specific methylation of distinct histone tails or lysine residues, or required for mono- or di- methylation rather than the H3K27me^3 mark.

There are few known Pc-G targets and their ectopic expression in *swn- clf-* that cannot solely account for the dramatic and pleiotropic phenotype. In addition, there are ~4400 genes possessing H3K27me^3 in the *Arabidopsis* genome that are likely Pc-G targets suggesting that the Pc-G plays a more central role than was previously thought to repress expression of key developmental regulators. ChIP on chip data in *Drosophila* has proposed the Pc-G to have between 100-200 targets, these were suggested targets based on loci being decorated with the H3K27me^3 mark, which localized with Pc-G proteins. This indicates many more targets of the Pc-G in *Arabidopsis* remain to be identified.

Although there is no obvious homologous PRC1 complex in *Arabidopsis*, there may be functional equivalents capable of stably repressing Pc-G targets. Likely candidates include *TERMINAL FLOWER2* (*TFL2*), and *EMBRYONIC FLOWER1* (*EMF1*). *TFL2* is a homologue of the *HETEROCHROMATIN PROTEIN 1* (*HP1*) gene found in *Drosophila* and has been shown to remodel euchromatin into heterochromatin to co-ordinate stable gene repression, together with the *Drosophila* PRC1 (de la Cruz *et al.*, 2007). ChIP on chip of *TFL2* has been shown to co-localize with ~80% of genes

possessing H3K27me^3 mark on chromosome 4 in euchromatic regions (Turck *et al.*, 2007). *TFL2* is required to repress *FLC, MEA, CRABCRAW* (*CRC*), *AP3,* and *LFY* as their expression is elevated in *tfl2*- mutants (Kotake *et al.*, 2003; Nakahigashi *et al.*, 2005). As *FLC, MEA,* and *AP3* are known targets of the Pc-G, it provides strong evidence that *TFL2* is likely to have retained its ability to repress PRC2 targets independently of the PRC1. Although, the analysis of the chromatin status of *TFL2* targets has not been widely characterized there is strong evidence to suggest *TFL2* is a functional equivalent of PRC1 in *Arabidopsis.*

Another putative PRC1 is *EMF1,* which acts as a transcriptional regulator shown to interact with nucleosome extracts and required to maintain H3K27me^3 marks on genes encoding transcription factors including *STM, AG,* and *AP2* (Calonje *et al.*, 2008). The expression of these genes are increased in *emf1*- mutants and this correlated with loss of H3K27me^3 marks at their loci (Aubert *et al.*, 2001; Li *et al.*, 2007). *emf1*- plants show a similar yet more severe phenotype than that of the Pc-G member *emf2*-. This indicates that *EMF1* plays a role in repressing gene expression of key developmental factors through histone methylation, and this may be in association with the Pc-G as *STM, AG, AP2* are direct targets of the Pc-G.

1. 7. Few direct Pc-G targets of the Pc-G are known

Until this investigation, very few direct targets of *SWN* and *CLF* had been identified, they include: *STM, FLC, AP3, FUSCA3* (*FUS3*), *AGAMOUS LIKE19* (*AGL19*), *MEA,* and *PHE1*. These genes were initially identified as Pc-G targets on the basis of three criteria. (1) their expression was increased in Pc-G mutants, (2) a Pc-G member was found enriched on their loci, and (3) they possessed H3K27me^3 in wild-type plants (Chanvivattana *et al.*, 2004; Schubert *et al.*, 2006; Makarevich *et al.*, 2006; Schonrock *et al.*, 2006; Katz *et al.*, 2004; Kohler *et al.*, 2003b).

The finding that there are more than 4000 genes decorated by the H3K27me^3 mark in *Arabidopsis,* and that the majority of H3K27me^3 is lost in *swn- clf*- mutants indicates

that the Pc-G complex(es) are the main source of creating the H3K27me^3 mark in euchromatin, and therefore the Pc-G complexes are likely to have many targets. The *swn- clf-* phenotype suggests massive transcriptional mis-regulation, of both direct Pc-G targets and indirect targets but what these are and their functions are unknown. The analysis of the *swn- clf-* double mutant phenotype may provide insight into the potential roles for Pc-G in plant development.

This thesis aimed to dissect the unique and redundant roles of CLF and SWN in processes known to be regulated by the Pc-G and to identify novel roles in development. This knowledge is crucial to understand the underlying basis for *swn-*, *clf-* and *swn- clf* mutant phenotypes and of the Pc-G in general.

1 . 8. Potential epigenetic control of embryo maturation

Chromatin remodelling genes and the Pc-G has been implicated in regulating the seed maturation program. The maturation phase of embryogenesis is characterized by the accumulation of storage oils and seed storage proteins (SSPs), and cell expansion of the embryo (Wobus & Weber, 1999). This process in essential for plants because the seeds of mutants that lack SSPs, and storage oils are desiccation intolerant and fail to undergo dormancy, which aids in future success of the plant (Luerssen *et al.*, 1998; Parcy *et al.*, 1997; West *et al.*, 1994; Parcy *et al.*, 1997; Stone *et al.*, 2001). This process is regulated by so called "embryonic master regulators" such as *FUSCA3* (*FUS3*), *LEAFY COTYLEDONS2* (*LEC2*), *LEC1*, and *ABSCISSIC ACID INSENSITIVE 3* (*ABI3*). These genes are specifically expressed during embryogenesis and promote the expression of genes that metabolize the seed specific storage oils and proteins, like SSPs during early and mid phases of maturation, and late embryogenesis-abundant proteins (LEAs) which accumulate in late embryogenesis (West *et al.*, 1994; Stone *et al.*, 2001; Luerssen *et al.*, 1998; Parcy *et al.*, 1997).

The evidence that these genes are central to the maturation process is based on *fus3-*, *lec1-*, *lec2-*, and *abi3-* mutant embryos being deficient in embryonic storage proteins

(Luerssen *et al.*, 1998; Parcy *et al.*, 1997; West *et al.*, 1994; Parcy *et al.*, 1997; Stone *et al.*, 2001). In addition, the mis-expression of *LEC1* post-germination, for example 35s::*LEC1* transgenic plants, exhibit accumulation of seed storage proteins, and increased *FUS3* and *LEC2* expression following germination (Kagaya *et al.*, 2005) .This indicates that *LEC1* acts upstream of *LEC2* and *FUS3* and that it promotes seed storage accumulation. *ABI3*, *LEC2*, and *FUS3* encode B3 class transcription factors that act to promote gene expression via a binding motif found in the promoter sequences of SSPs and LEAs (Suzuki *et al.*, 2007; Wobus & Weber, 1999). The phyto-hormone ABA is known to positively regulate seed maturation, by promoting the expression of *ABI3*, *ABI4* and *ABI5*, which are required for correct embryo maturation by promoting storage compound accumulation (Carles *et al.*, 2002; Soderman *et al.*, 2000; Penfield *et al.*, 2006; Parcy *et al.*, 1997). *ABI4* encodes an *AP2*-like transcription factor, is known to play a role in maturation stage of embryogenesis (Penfield *et al.*, 2006; Soderman *et al.*, 2000) and ABI5 promotes the expression of seed storage proteins and physically interacts with ABI3 (Lopez-Molina *et al.*, 2002).

MEA is found enriched at the *FUS3* loci and the H3K27me^3 found at the *FUS3* locus is reduced in developing *mea*- seeds (Makarevich *et al.*, 2006). *mea*- mutants show elevated *FUS3* expression, as do the direct targets of *FUS3* including SSPs and LEAs (Kohler *et al.*, 2005; Makarevich *et al.*, 2006). This suggests that the *MEA*-Pc-G acts to repress the maturation master regulator *FUS3* in the endosperm of the seed where it is not normally repressed, as demonstrated in reporter gene studies (Makarevich *et al.*, 2006).

In conclusion, the regulation of the seed maturation program is largely at the mercy of the master regulators (genetic regulation) and hormonal processes, and is also regulated by the histone methylation by the action of the Pc-G. A similar multi-regulatory pattern is seen in the repression of maturation traits post germination.

1. 9. The Pc-G may repress seed maturation specific expression in post germinative development

Although the accumulation of storage proteins and oils is essential during embryo maturation, it is equally important these embryonic traits are not expressed during post-germinative development, which is demonstrated in a series of mutants described below. There are several known mechanisms of repressing embryonic traits post germination, including chromatin remodelling, histone deacetylation, and the use of repressive B3 class transcription factors.

PICKLE (*PKL*) encodes a CHD3 ATP dependent chromatin remodeller, which when mutated show accumulation of both storage oils and proteins and a "pickle root" phenotype that resembles a stunted opaque lateral root (Dean Rider S Jr *et al.*, 2003; Li *et al.*, 2005; Ogas *et al.*, 1997; Henderson *et al.*, 2004). The pickled root phenotype is enhanced when mutants are grown on GA inhibitor (uniconazole), illustrating that *PKL* represses embryonic traits through chromatin remodelling and this repression is partially GA-dependent (Henderson *et al.*, 2004; Ogas *et al.*, 1997). Another chromatin remodeller, *BRAHMA* (*BRM*), also shows a role in repressing embryonic storage traits as *brm*- mutants accumulate embryonic proteins in its leaves (Tang *et al.*, 2008). These results indicate that the *PKL* and *BRM* chromatin remodellers may have similar roles to repress embryonic traits in the different tissue types.

HIGH SUGAR EXPRESSION LEVEL1 (*HSL1)* and *HIGH SUGAR INDUCIBLE2* (*HSI2)* encode B3 class transcription factors related to *FUS3, LEC1, LEC2* (Tsukagoshi *et al.*, 2007). The *hsl1- hsl2-* double mutants exhibit ectopic expression of seed storage proteins and oils post-germination when grown on high sucrose media. The mutants also show phenotypic deformities including stunted growth and ectopic cell division on the root epidermis before dying (Tsukagoshi *et al.*, 2007). *HSL1* and *HSI2* are thought to act redundantly to repress the effects of sugar to promote ectopic expression of embryonic traits. Sugar, in the form of glucose or sucrose, is critical for the correct regulation of embryo maturation as it can induce the expression of *ABI3, ABI4, ABI5* and LEAs

independently of ABA in post-germination development (Rook *et al.*, 2006), a transcription profile reminiscent of the early stages in maturation.

swn- clf- mutants often show a *pkl-* like "pickle" root phenotype, and *FUS3* is found up regulated in both *pkl-* and *swn- clf-* double mutants. CLF is found enriched at the *FUS3* locus in wild-type embryos and seedlings respectively (Makarevich *et al.*, 2006). *FUS3* is not up regulated in *clf-* or *swn-* plants suggesting it is a redundant target (Makarevich *et al.*, 2006). In addition, the morphological defects of *pkl-* and *hsi1- hsl2-* are similar to the *swn- clf-* double mutant phenotype. Together the evidence highlights a redundant role of *SWN/CLF*-Pc-G may be to repress embryonic traits during post-germinative development, this hypothesis is later explored.

In summary, there is some evidence to suggest that the Pc-G may regulate embryogenesis seed maturation and ensures its confinement to seed development, but it remains unclear if the similarities were superficial or due to the same mechanistic causes.

1. 10. Concluding remarks and pertinent questions

The role of *SWN* in plants is particularly interesting as it enhances both *clf-* and *mea-* mutant phenotypes yet shows no obvious defects, illustrating that it is of fundamental importance in development but largely acting redundantly. This raises the question: Does *SWN* have roles in development that are independent of *CLF* and *MEA*? Uncovering the discrete and redundant roles of *SWN* in development and identification of its targets will add to our understanding of the Pc-G in plants. There are also points of evidence to suggest that *SWN* may discrete roles in development: Firstly, *SWN* and *CLF* homologues are found in a wide range of plant species including *Arabidopsis*, petunia, maize, and rice (Mayama *et al.*, 2003). This demonstrates that *SWN* and *CLF* have been conserved since the ancient divergence of monocots and dicots. As both *SWN* and *CLF* have not been lost during evolution, it strongly suggests they have important roles in development independent of each other and *MEA*. For example, if *SWN* only acts

redundantly with *CLF* and *MEA* in development there would have been no selective pressure to conserve *SWN*.

Secondly, the role of the Pc-G in vernalization is known, but *clf*- plants are vernalization responsive. It is not yet known if null *swn*- plants respond to vernalization. In addition to this, there are definite similarities between phenotypes of *swn*- and *clf*- plants and the other Pc-G members *vrn2*- and *emf2*- mutants. For example, the *vrn2*- mutants are not sensitive to vernalization, has no gross morphological defects, whereas *emf2*- plants are early flowering with curled leaves. Both phenotypes are reminiscent of *swn*- and *clf*-, respectively, apart from the vernalization phenotype of *vrn2*-. This is suggests that *SWN*, like *VRN2*, is involved discretely in the vernalization response.

Thirdly, the expression patterns of *SWN* and *CLF* are marginally different, *SWN* expression is higher than *CLF* in the siliques and in response to stress, suggesting that *SWN* may play a role independent of *CLF* in these respects. Although circumstantial, each of the known Pc-G members that are expressed post-germination all show flowering and J-A transition defects, illustrating that the Pc-G is essential in regulating phase transitions. This may suggest that subtle phase transition defects will be observed in *swn*-.

Finally, internet based microarray database (Genevestigator) indicates that *SWN* and *CLF* have slightly different expression patterns during ontogeny, in response to ABA and stress treatments of wild-type seedlings. This suggests that *SWN* has acquired a discrete role in development. These relative differences in *SWN* and *CLF* expression could help to isolate aspects of development that are controlled by *SWN*, independently of *CLF* and *MEA*. Careful observation of the response of *swn*- mutants to a variety of treatments, such as the response to vernalization, hormones, and flowering time may reveal SWN functions.

The *swn- clf*- phenotype is incredibly severe but the causes of this dramatic phenotype are entirely unknown. I plan to investigate the causes of the root phenotype, in

particular, and explore the other phenotypic deformities of the *swn- clf-* double mutant phenotype to identify novel aspects of development that are regulated by the Pc-G complex. This will be achieved through phenotypic characterization and transcriptional profiling of *swn-, clf-* and *swn- clf-* mutants. Transcriptional profiling of these mutants will also allow the resolution of the proposed partial functional redundancy between *SWN* and *CLF* and potentially identify novel Pc-G targets on the genome scale. This may also reveal a "snap shot" the fundamental roles of the Pc-G in regulating gene expression controlling development at least at one development stage.

The Pc-G complex obviously plays many important roles in development by repressing gene expression, yet few direct targets are known. The *swn- clf-* phenotype and mis-expression of *STM* and *FUS3* has suggested a role for the Pc-G in maintaining stem cell identity and repression of embryonic traits post-germination. The full extent of just how many targets the Pc-G has is unclear but given the phenotype of Pc-G mutants, a large number remain to be identified. Analysis of the early *swn- clf-* may provide evidence of further roles of the Pc-G in development. Understanding the role of the Pc-G in development and identifying its direct targets are central to the investigations in this thesis.

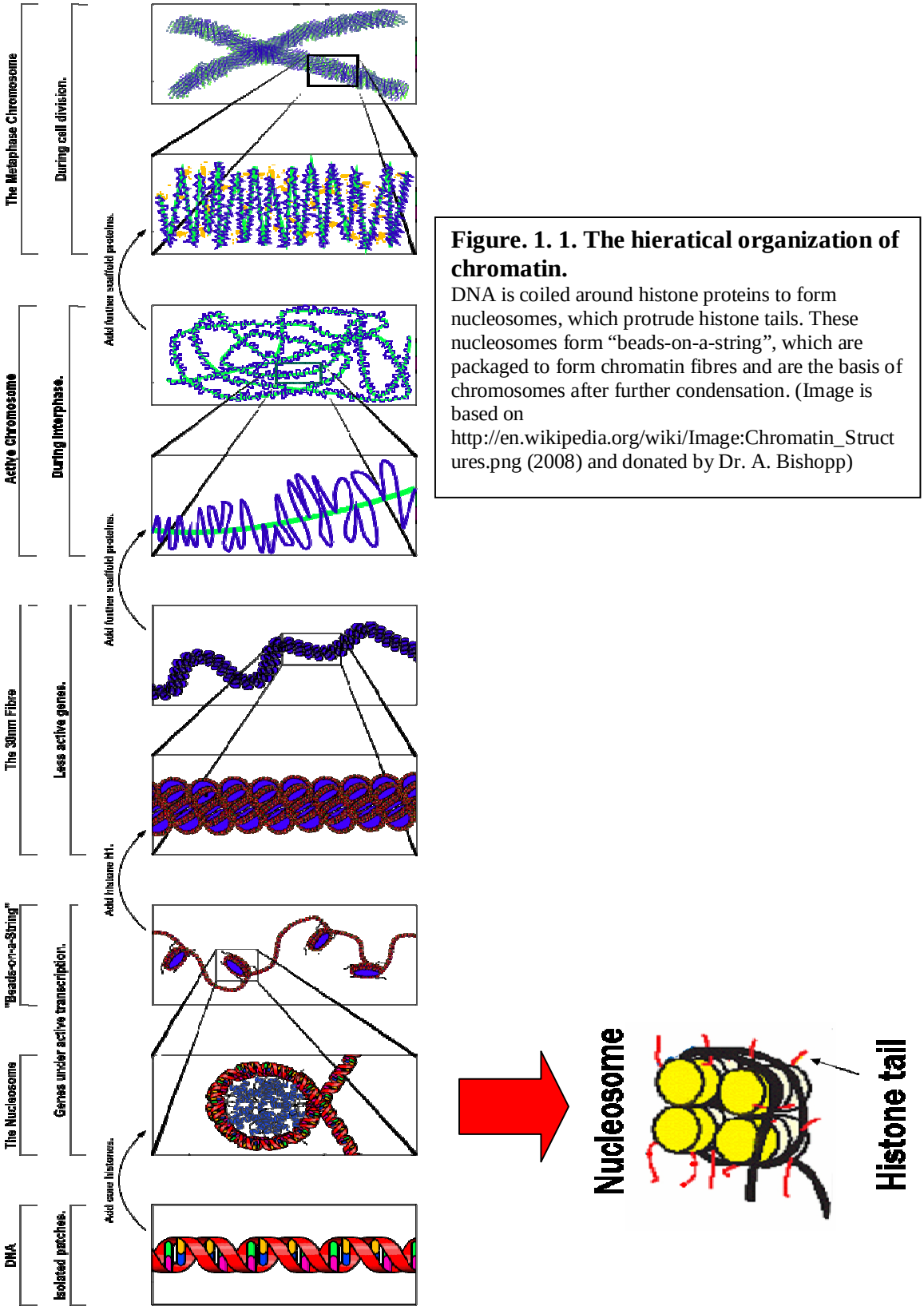

Figure. 1. 1. The hieratical organization of chromatin.
DNA is coiled around histone proteins to form nucleosomes, which protrude histone tails. These nucleosomes form "beads-on-a-string", which are packaged to form chromatin fibres and are the basis of chromosomes after further condensation. (Image is based on http://en.wikipedia.org/wiki/Image:Chromatin_Structures.png (2008) and donated by Dr. A. Bishopp)

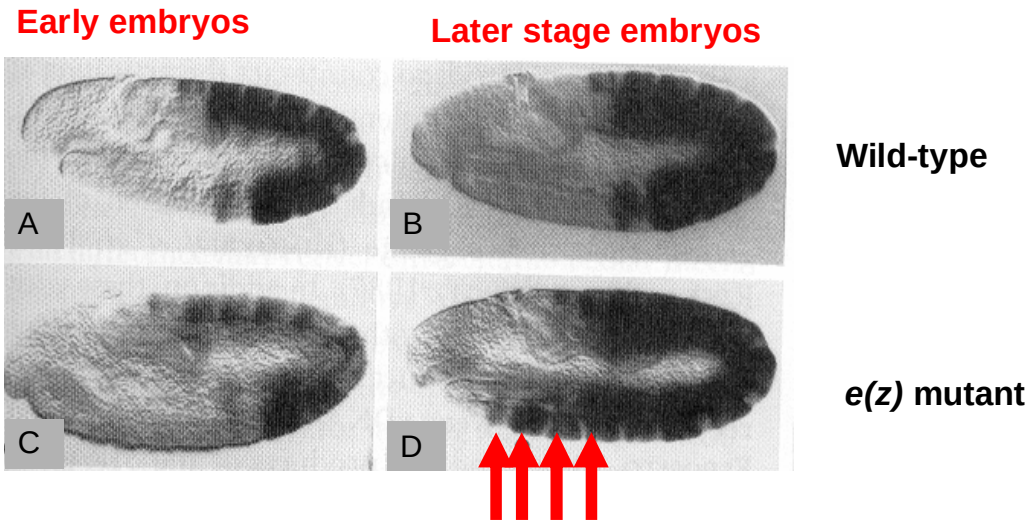

Fig. 1. 2. *UBX* expression in wild-type and *E(z)*mutant *Drosophila* embryos.
In wild type early larval stages (a) *UBX* expression is confined to the posterior region and is maintained there in late larval development (e). In the early stages of *E(z)* mutant larvae *UBX* is also confined to the posterior segments (b), however in later stages of larval development *UBX* if ectopically expressed in anterior and posterior regions (f), as indicated by arrows. (*In situ* hybridization of *UBX*, image taken from (Jones & Gelbart, 1990))

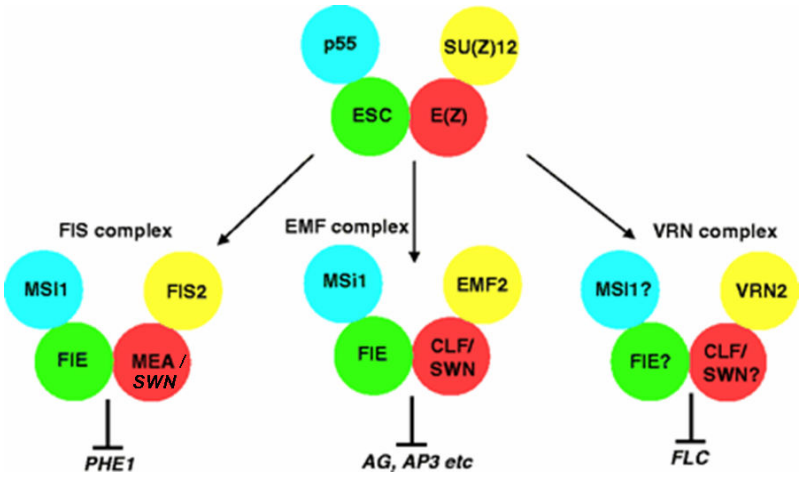

Fig. 1. 3. Different Pc-G complexes regulate different aspects of development.
"The core components of the *Drosophila* PRC2 complex are shown at top. In *Arabidopsis*, an equivalent ancestral complex is proposed to have diversified into three similar complexes with at least partially discrete functions. The colours indicate homology; so for example, E(z) homologues are coloured red. The contacts indicate interactions; for example, FIE can interact with MEA and MSI1 but not FIS2, whereas FIS2 can interact with MEA but not with other FIS proteins. The target genes shown are not comprehensive; it is likely that all three complexes have many more targets than those shown" (Chanvivattana *et al.*, 2004). The FIS complex represses endosperm proliferation, the EMF2 complex

regulates flowering time and flower identity, and the VRN2 complex regulates vernalization by repressing *FLC*.

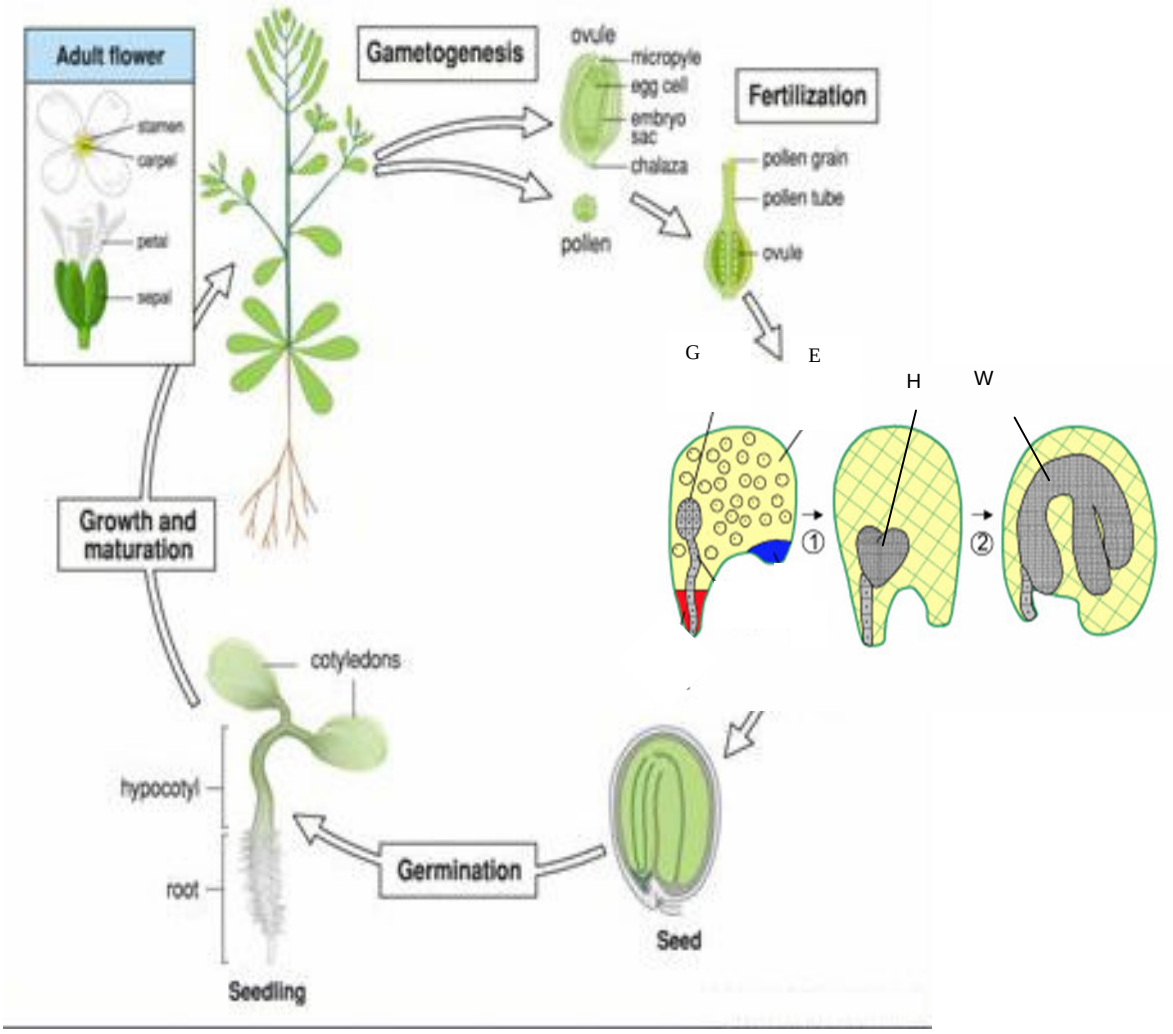

Fig. 1. 4. Life cycle of *Arabidopsis thaliana*.
Following double fertilization, the central cell divides and fills the ovule before undergoing cellularization (E) and it ultimately consumed by the embryo (absence is denoted by hatching). The egg cell divides to form the embryo, and goes through the globular (G), heart (H), torpedo, and the walking stick stages (W). The seed then undergoes metabolic maturation and is subsequently dormant. Given the correct environmental cues germination occurs. The seedling grows and goes through the juvenile to adult stages of development and is then competent to flower given both endogenous and environmental cues. (Johnston *et al.*, 2007) and http://phy.asu.edu/phy598-bio/D7%20Notes%2006%20Part%202_files/image027.jpg (2008).

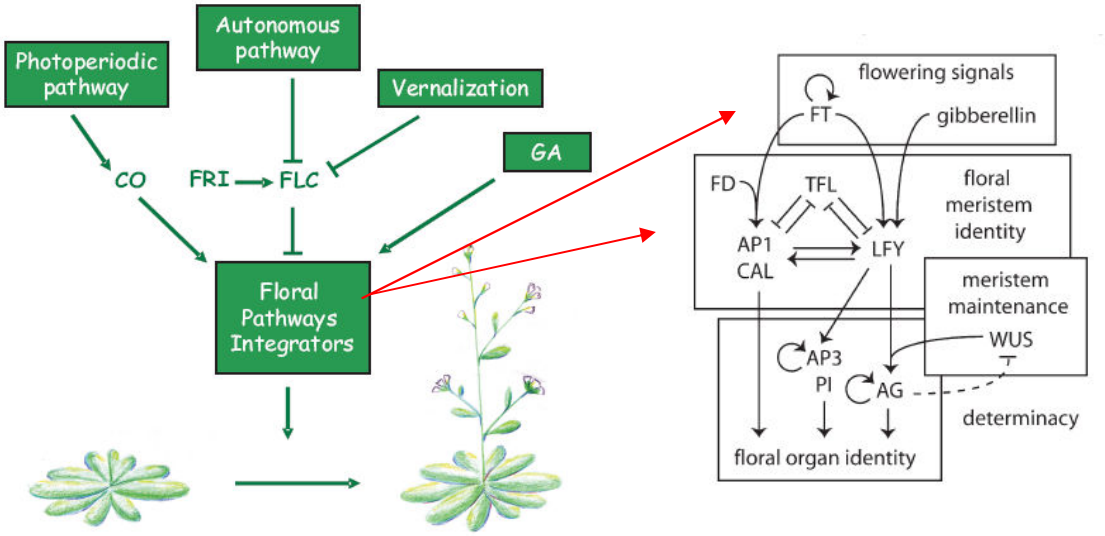

Fig. 1. 5. The genetic network regulating flowering.
Diagram of the signals effecting the flowering transition, floral meristem and organ identity. The flowering pathways involve a series of genetic steps for flowering and flower formation to occur. Note the promotion of *FLC* by *FRI*, the repression of *FLC* by vernalization, and that *FLC* represses the floral meristem identity genes. (Sablowski, 2007; Parcy, 2005).

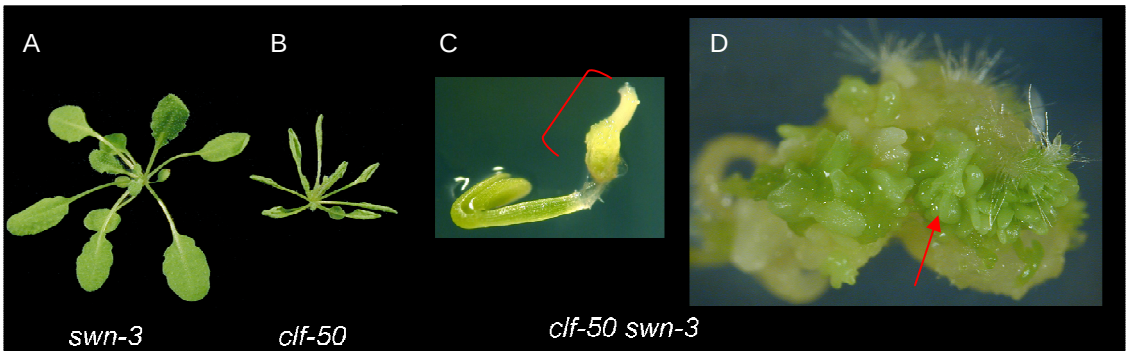

Fig. 1. 6. Post-germinative phenotypes of *Arabidopsis* Polycomb mutants.
swn-3 plants (and all other *swn-* alleles) are morphologically indistinguishable from wild-type plants (A). *clf-50* mutants (other *clf-* alleles show varying degrees of phenotypic severity) are small, early flowering, have epinastic leaf curling, and reduced fertility due to partial homeotic transformation of flower tissues (B). The *swn-3 clf-50* mutants show a “pickle” root phenotype (bracket) (C), fail to form any true organs and continue to proliferate, developing into the callus-like tissue (D). Occasional somatic embryos are observed on the callus-like tissue, and organ-like projections (arrowed) form the basis of the callus-like tissue (Chanvivattana *et al.*, 2004).

1. 11. Aims

The overall aims of this thesis are therefore to:

- Discover if *SWN* plays a non-redundant role in development in phase transitions, stress, and hormonal response.
- Uncover which aspects of post-germinative development are *SWN*-*CLF*-Pc-G regulated.
- Identify potential targets of *SWN* and *CLF*.

2. 0. Developmental analysis of *swn*- mutants

As outlined in the introduction, several lines of evidence lead me to investigate whether *SWN* has a function independent of either *CLF* or *MEA* in development. As *swn*- plants have no gross morphological defects under normal growth conditions, a detailed analysis was carried out comparing wild-type and *swn*- mutants with respect to phase transitions, and growth in specific conditions in which expression of *SWN* and *CLF* differ.

2. 1. Do *swn*- seedlings respond differently to hormonal treatments or high salinity?

Phyto-hormones are very well known regulators of plant development. There are eight main phyto-hormones including gibberellic acid (GA), abscisic acid (ABA), auxin, jasmonic acid, salicylic acid, brassinosteroid, ethylene, and cytokinin. Each of these compounds has been shown to play important roles in development (functions are reviewed in (Gubler *et al.*, 2005; Debeaujon & Koornneef, 2000; Okamuro *et al.*, 1996; Jenik & Barton, 2005; Nemhauser *et al.*, 1998; Swarup *et al.*, 2002; Hartig & Beck, 2006; Gray, 2004; Halim *et al.*, 2006)).

ABA acts in seeds to inhibit germination, or induce dormancy, but it also has a role in stress response and stomata aperture regulation. Mutants devoid of ABA, or in the signalling of ABA, germinate earlier than wild-type, and are stunted in growth (Gubler *et al.*, 2005; Koornneef *et al.*, 1989). GA is considered antagonistic to ABA, as it is involved in the promotion of germination, or breaking seed dormancy. Plants lacking GA have very low germination rates. Mature plants lacking GA are also small with reduced fertility (Debeaujon & Koornneef, 2000). Phenotypes of mutants that cannot synthesize hormones can be reverted when the missing hormones are exogenously applied. The application of ABA to wild-type seedlings show reduced germination, and shorter root growth depending on concentration (Koornneef *et al.*, 1989). Conversely, GA promotes germination and increases hypocotyl length (Debeaujon & Koornneef, 2000). Internet based microarray expression data (Zimmermann *et al.*, 2004) showed that when seedlings were exposed to applied phyto-hormones, and salinity, there was a

vast difference in global expression, compared to those without treatment. When wild-type seedlings are treated with ABA, *SWN* expression is relatively higher than *CLF* (Fig. 2. 1). A similar result was observed when wild type seedlings were grown in stressful salt conditions (Zimmermann *et al.*, 2004).

To test whether *SWN* plays a role in ABA response *swn*- and wild-type plants were grown on ABA containing media. Germination rate, root and hypocotyl growth were assayed as a measure for response to ABA. Seedlings were scored every 3 days for germination, root length and hypocotyl length. No difference was detected in germination rate, root or hypocotyl growth between *swn*- and wild-type plants after 15 days. Known ABA growth responses were observed, such as reduced germination rate and root growth, but *swn*- plants did not differ from wild-type (data not shown) (Koornneef *et al.*, 1989; Gubler *et al.*, 2005). As GA acts as an antagonist to ABA, growth of *swn*- and wild-type plants were assayed on media containing GA. There was no difference in germination, root, or hypocotyl length of *swn-3* wild-type plants grown on GA media (data not shown).

Salt stress was another condition where *SWN* expression was relatively higher than *CLF* in microarray experiments (Zimmermann *et al.*, 2004). Measuring root and hypocotyl length, no difference was detected between wild-type and *swn*- plants when grown on high salt conditions (Table. 2. 1). The expected salt stress phenotype (slow and reduced growth) (Peng *et al.*, 2007) was observed in both genotypes.

Mutants in auxin signalling and regulation show wide ranging developmental defects, including embryo lethality, aerial parts lacking organs (e.g. flowers), altered leaf shape, and reduction in root hairs (Hunter *et al.*, 2006; Jenik & Barton, 2005; Leyser, 2005). Wild-type seedlings supplied with auxin show increased root growth, relative to controls grown without cytokinin (Sun *et al.*, 2003)

Cytokinin regulates cell division, senescence, lateral bud formation, and chloroplast maturation. Wild-type roots supplied with cytokinin show stunted root growth and an increased lateral root formation (Woo *et al.*, 2007).

The effects of auxin and cytokinin on *swn*- plants were also assayed because the interaction of hormones has also been shown to be important in cell division and expansion in the meristems (Shani *et al.*, 2006), so growth may have been affected in *swn*- plants. No difference was observed between wild-type and *swn*- root and hypocotyl length, when grown on auxin or cytokinin media. The expected phenotypes were observed for exogenous application: for auxin, reduction in primary root growth; and increase in hypocotyl length (Woo *et al.*, 2007), and for cytokinin increased root growth and reduced hypocotyl growth (data not shown) (Sun *et al.*, 2003).

In conclusion, it appears that *SWN* is not acting discretely to regulate the response to applied ABA, GA, auxin or cytokinin in terms of root or hypocotyl growth or germination. In addition, I was unable to find any function for *SWN* in the response to salt stress in root or hypocotyl growth or germination. The lack of phenotypic differences in response to phyto-hormones or salt stress could be explained by *CLF* masking the effect of losing *SWN*.

As no suggestion of a discrete *SWN* function came from this approach, a different strategy was taken, by looking into developmental transitions. The aim was to discover whether *SWN* plays a non-redundant role in regulating phase changes.

2. 2. Are phase transitions affected in *swn*- plants?

Arabidopsis undergoes several dramatic changes in growth phase in its life cycle, including embryogenesis, dormancy, juvenile vegetative phase, adult phase, flowering, and senscencing. Each of these phases of life are strictly genetically regulated and coordinated by environmental cues (Komeda, 2004; Poethig *et al.*, 2006; Lim *et al.*, 2007).The juvenile to adult transition (J-A transition) and vegetative to flowering

transition (V-F transition) are of most interest with regard to potential regulation by *SWN*. Reducing expression of Pc-G members tends to cause early flowering. *clf-*, *emf2-*, and co-suppressed *FIE* show varying degrees of early flowering (Goodrich *et al.*, 1997; Chen *et al.*, 1997; Kinoshita *et al.*, 2001). However, *FIS2* and *MEA* do not affect flowering because their expression is restricted to seed development (Luo *et al.*, 2000). This shows that one of the functions of the Pc-G complex is to repress flowering, and raises the possibility that *SWN* may function non-redundantly in controlling flowering time.

The phenotype of *emf2-* mutants is particularly interesting with respect to phase changes, as *emf2-3* mutants undergo the J-A transition immediately, they fail to form juvenile leaves, and flower ~one week after germinating (Chen *et al.*, 1997). This suggests that *EMF2* acts as a strong floral repressor, or that it represses the J-A transition. It is possible that *EMF2*, and therefore the Pc-G complex, is involved in juvenile leaf production, J-A transition and the V-A transition.

Both the J-A and V-F transitions were analyzed to identify a possible discrete role for *SWN* in development. Although no gross phenotype of *swn-* plants has been observed, this was one of the most likely places to uncover a phenotype. To investigate a possible role for *SWN* in flowering, wild-type and independent *swn-* alleles were sown and their flowering times recorded and compared. The hypothesis was that *SWN* might function in the same manner as other Pc-G members to repress flowering, so the *swn-* plants may have subtle effects on flowering time.

2. 2. 1. Does *SWN* play a role in regulating flowering time?

As can be seen in figure 2. 2 there was no difference in the number of days to flower between *swn-* mutants and wild-type progenitor plants in long days. As an alternative measure of flowering time, the number of rosette leaves at the time of one cm bolt was recorded for wild-type and *swn-* alleles. This can illustrate possible defects in leaf initiation rate. *swn-* lines produced significantly more leaves than wild-type in long days

(Fig. 2. 3). However, *swn-* lines and wild-type flowered after the same number of days, illustrating that *swn-* lines have an increased leaf initiation rate. Therefore, *SWN* is likely to act to repress leaf initiation rate rather than by regulating flowering directly.

The effects of *swn-* mutations on leaf initiation rate, days to flower, and number of leaves at flowering were also observed when grown in long days at 30°C (appendix: Fig. 8. 1, Fig. 8. 2).

The effects of *swn-* mutations on flowering time and leaf initiation rate were tested in short day conditions, as the endogenous signals and environmental cues required for flowering in short and long days are different (Imaizumi & Kay, 2006). The effect of *swn-* on flowering in short days could therefore be different to that in long days and subtle differences are often more apparent under short days because it requires more time to flower. Interestingly, *clf-* plants have a slightly suppressed or relieved phenotype when grown under short day conditions (personal communication, Dr. W. J. Goodrich). This suggests some interaction of day length and Pc-G function. In short days it took, on average, wild-type 105 days to flower, *swn-2* 79 days and *swn-7* 72 days, however, it took *swn-3* mutants 219 days to flower (Fig. 2. 4). It is therefore difficult to evaluate the effect of *swn-* mutations on flowering time, as there is no consistent effect of *swn-* on flowering time. As there is a conflict of phenotype of mutations in the same gene it is possible that the flowering time phenotype of the *swn-3* allele is due to dominant effect of the mutation causes a delay in flowering, the *swn-2* and *swn-7* mutant alleles show a earlier flowering phenotype compared to wild-type plants. This is consistent with the possibility that *SWN* does indeed play a role in repressing flowering, but further work would be required to confirm this.

Figure 2. 5 shows that *swn-* lines have, on average, slightly more leaves than wild-type plants, and *swn-2* and *swn-7* flowered earlier than wild-type plants. This suggests that the leaf initiation rate is once again altered. Figure 2. 6 demonstrates that each *swn-* mutant line has a higher leaf initiation rate in short days, as is the case in long days.

The role of *SWN* in flowering in short days remains ambiguous, however, it is likely that *SWN* plays a role in repressing leaf initiation rate in short and long days. This indicates that *SWN* regulates flowering time through the autonomous pathway as a similar effect is observed in long and short days.

2. 2. 2. Does SWN play a role in the vegetative phase transition?

The J-A transition is judged as the time in which the first adult leaf is formed (Telfer *et al.*, 1997). The evidence that *swn*- mutants have an increased leaf initiation rate raises the question of whether they are defective in the J-A transition. This is because a number of mutants which have an increased leaf initiation rate have an altered J-A transition, as for example is observed in *HASTY* (*hst*-) (Telfer & Poethig, 1998) and *SERRATE* (*ser*-) mutants (Clarke *et al.*, 1999). However, the change in leaf initiation rate does not always changes the timing of the J-A transition, as for example is shown in the *PAUSED* (*psd*-) mutant which has a normal J-A transition in time but form few, if any, juvenile leaves (Hunter *et al.*, 2003).

To investigate whether *SWN* functions in the J-A transition, wild-type and *swn*- mutants were scored for number of juvenile leaves, and the time of the J-A transition. The J-A transition is the time it takes for the first adult leaf to be produced, which unlike juvenile leaves form abaxial trichomes.

It was found that in both long and short days independent *swn*- lines possessed more juvenile leaves than wild-type plants (Fig. 2. 7 and Fig. 2. 8). However, as *swn*- plants have an accelerated leaf initiation rate, more juvenile leaves would be predicted, and so the J-A transition may remain unaffected. It was found that all independent *swn*- lines did undergo the phase change later than wild-type plants in both long and short days (Table. 2. 2).

This shows that not only do *swn*- mutants produce leaves faster than wild-type plants, but that they are delayed in the J-A transition in long and short days. The effect of *swn*-

on the J-A transition is enhanced in short day conditions, this is shown by the J-A transition delay in *swn*- lines compared to wild-type plants being larger in short days than long days (Fig. 2. 2).The flowering time and juvenile leaf number experiments were repeated twice and similar results were obtained (data not shown).

In conclusion, in long and short day conditions *swn*- plants show increased leaf initiation rate compared to wild-type plants. This could explain why *swn*- plants flower with more leaves than wild-type, and possess more juvenile leaves. It is known that the rate of leaf production is not necessarily correlated with the J-A transition (Telfer *et al.*, 1997). Although many mutants that have an altered leaf initiation rate also have an altered J-A transition, these mutants frequently show flowering time defects (e.g. *ser*- mutants (Clarke *et al.*, 1999)). However, *swn*- mutants exhibit no flowing time defects. Therefore, I suggest *SWN* acts non-redundantly to regulate both the J-A transition and leaf initiation rate.

2. 2. 3. Analysis of the co-segregation of the delayed J-A transition in *swn*- plants

Independent *swn*- lines showed subtle effects on the J-A transition suggesting that *SWN* regulates the transition. To confirm that the *SWN* mutations were responsible, rather than linked mutations or stochastic variation, *swn*- mutants were backcrossed to *SWN+* progenitor line (Col-0) and the resulting F2 populations were analyzed. The strongest phenotypic effect was found in *swn-3/-3* so this allele was used for testing co-segregation.

F2 plants were scored for juvenile leaves, flowering time and total number of leaves. Plants were considered as "mutant" when they possessed ≥10 juvenile leaves, as wild-type plants grown simultaneously showed a maximum of 9 juvenile leaves. The J-A transition appears to be a variable process, because in previous experiments wild-type plants had ~5 juvenile leaves and *swn*- lines had 6 -10 juvenile leaves in long days (Fig. 2. 7). Those plants judged to have a "mutant" phenotype were genotyped. If the *swn-3* mutation was responsible for the phenotype then all of the genotyped plants should have

been *swn-3/-3*. If *swn-3* mutation were not responsible, one would expect only 25 % of the genotyped plants to be homozygous due to independent segregation of the *swn-3* allele.

In long days, 78 plants were scored of which 25 plants were judged to have a "mutant" phenotype (10 ≥ juvenile leaves) and were genotyped. Of the 25 genotyped 13 plants were found to be *swn-3/-3* (52%). A χ^2 test was used to assess if the *swn-3/-3* genotype occurred significantly more than expected. This shows that *swn-3/-3* is indeed found significantly more than expected (Table. 2. 3 ($P < 0.01$)). *swn-3/+* plants also showed a significantly higher frequency than expected. A similar result was obtained in short day conditions, however, in this case 60% (9/15 plants) of the plants genotyped were *swn-3/-3*, (Table. 2. 4). Indeed *swn-3/-3* is found more frequently than expected if *swn-3/-3* was not responsible for the increased number juvenile leaves (χ^2 analysis $P < 0.01$). Therefore, the *swn-3/-3* mutation can be said to cause a tendency to increase the number of juvenile leaves compared to wild-type, but as it is unable to consistently show the mutant phenotype may suggest that it is an epi-mutant or epi-allele and its phenotypic effects are unstable.

The failure to observe complete co-segregation may have occurred for two possible reasons: 1) A linked mutation to *swn-3*, although this is unlikely as independent *swn-* lines showed increased numbers of juvenile leaves. 2) It is possible that *swn-* mutation was not fully penetrant and stochastic variation caused the increased numbers of juvenile leaves.

To test these two potential causes of phenotypic variation I scored F3 populations from F2 homozygous individuals. The average number of juvenile leaves of 12 plants (in 11 F3 families) were scored to reduce the effects of stochastic variation and increase the effect of heritability. The eleven families of *swn-3/-3* and *SWN+/+* were selected randomly after genotyping F2 plants.

Progeny of eleven homozygous F2 plants of *SWN+* and *swn-3/-3* genotypes were sown in long days. All eleven *swn-3/-3* families were found to possess more juvenile leaves than the *SWN+* families (Fig. 2. 9). The average of the eleven *swn-3/-3* families have significantly more juvenile leaves than the eleven *SWN+* families (as analyzed by Students t-test ($P < 0.01$)) suggesting *swn-3/-3* is responsible for the phenotype.

The fact that so many of the *swn-3/-3* families have more juvenile leaves, may suggest that the *swn-3* mutation is responsible for the juvenile leaf phenotype. However, if a secondary mutation is closely linked to *swn-3* then segregation may not necessarily be seen, and plants homozygous for *swn-3/-3* could be homozygous for the secondary mutation too.

No correlation was found between average numbers of juvenile leaves in the F3 families with its F2 parent in *SWN+* or *swn-3/-3* i.e. F2 *SWN+* plants that had a large number of juvenile leaves did not consistently show high number of juvenile leaves in the F3 generation. Conversely, *swn-3/-3* plants that showed "wild-type" number of juvenile leaves in the F2 generation often showed an increased number of juvenile leaves in the F3 generation. Therefore, the potential effects of secondary mutations causing the phenotype is not heritable, so the variation of the juvenile leaves phenotype is likely to be a consequence of environment influence. Although all the traits measured are prone to variation but the *swn-* phenotypes have been observed and are reproducible, in long and short days, in independent *swn-* mutant alleles, in blind co-segregation experiments, and in families originating from a single parent. In addition, each *swn-* line tested showed a tendency to possess an increased number of juvenile leaves, increased leaf initiation rate and delayed J-A transition.

It is notable that no significant difference in number of days to flower or percentage of juvenile leaves between the 11 families of *swn-3/-3* and *SWN+* was seen in long day conditions (appendix: Fig. 8. 2). 6/11 *swn-3/-3* families flowered with significantly more leaves than any of the *SWN+* families (Fig. 2. 10).

The *swn*- phenotypes display increased variation compared to wild-type. However, wild-type plants have shown "mutant" phenotypes. I suggest the variation in the number of juvenile leaves in *SWN*+ and *swn-3/-3* lines is due to environmental variation, and/or stochastic variation, rather than secondary linked mutations. As in multiple *swn*- individuals of the same genotype consistent differences are seen between *swn*- and wild-type.

To validate the possible role of *SWN* functioning in the J-A transition beyond doubt, a complementation experiment should be carried out. This could be achieved by introducing a genomic copy of *SWN* into a *swn-/-* background, and score the progeny, and testing whether plants carrying the *SWN* transgene have a "wild-type" phenotype i.e. a normal J-A transition. This was undertaken but the SWN transgene failed to express *in planta*, and time limitation made it impossible to pursue further.

Although the phase change phenotype of plants is extremely variable, *SWN* is likely to play a non-redundant role regulating the leaf initiation rate, the number of juvenile leaves, and the timing of vegetative phase change.

2. 3. *SWN* does not play a discrete role in the vernalization response.

The vernalization response was tested as vernalization is regulated by the Pc-G complex (outlined in the introduction). *clf*- mutants in a vernalization requiring background have been found to be vernalization sensitive (Wood *et al.*, 2006; Chanvivattana *et al.*, 2004). This suggests a possible discrete role of *SWN* acting in the Pc-G complex to regulate the vernalization response.

To discover whether *swn*- plants were responsive to vernalization *FRI*+ and *swn-3* plants were crossed, and the F2 generation was genotyped to identify individuals that gave rose to F3 families that were *FRI+/+ SWN+/+* and *FRI+/+ swn-3/-3*. Following vernalization treatment *FRI+/+ swn-3/-3* plants flower at the same time as *FRI+/+SWN+* plants in both long and short days (Fig. 2. 11 and data not shown). Each

genotype flowered with ~12 leaves, *FRI+/+* plants without the vernalization treatment flower with >100 leaves (Fig. 2. 12) showing that the vernalization treatment was effective. This illustrates that *SWN* does not have a discrete function in the vernalization response; instead, it seems likely that *SWN* and *CLF* are acting redundantly to regulate the vernalization response, by directly acting on the *FLC* locus.

When F3 seeds from *FRI+/+ swn-3/+* plants were grown without vernalization, plants that were *FRI+/+ swn-3/-3* flowered later than *FRI+/+ SWN+/+* and *FRI+/+ swn-3/+* showed an intermediate flowering phenotype (Fig. 2. 12). This suggests that *SWN* regulates flowering through the autonomous pathway with dependency on *FRI*. This theory is further examined in the discussion chapter.

2. 4. Summary and conclusions

Microarray data indicated *SWN* had relatively higher expression compared to *CLF* in response to ABA application, and salt stress. However, investigations aimed at uncovering a non-redundant role for *SWN* in these conditions were inconclusive. No morphological differences between *swn-* and wild-type plants were observed in terms of seedling growth or germination rate under any conditions tested.

Careful analysis of phase transitions in *swn-* plants provides evidence of subtle *swn-* mutant phenotypes. In both long and short day growth conditions *swn-* lines show a tendency to have increased leaf initiation rate, delayed J-A transition, and more leaves at the time of flowering than wild type plants.

To verify that *SWN* was responsible for the phenotype a F2 population segregating for *swn-3* was analyzed. This showed that the leaf initiation rate and the number of juvenile leaves phenotypes occurred predominantly but not exclusively with the *swn-3/-3* genotype. This was also supported by the scoring of eleven F3 families of *SWN+* and *swn-3/-3* families from single F2 parents. This showed that *swn-3/-3* families predominantly, but not exclusively, had significantly more leaves at flowering, an

increased juvenile leaves and an increased rate of leaf production in long days. The *swn*- mutant phenotypes are subtle and would likely have been missed in previous analysis.

It can be suggested that *SWN* is involved in the regulation of leaf initiation and the J-A transition. Whether this occurs as two separate processes, or the phase change delay causing the increased leaf initiation rate remains unclear. Theoretical mechanisms of how *SWN* may regulate these processes are discussed in a later chapter.

The third point of investigation was vernalization. It was found that *swn*- plants were fully responsive to vernalization, and so *SWN* does not play a discrete role in the vernalization response. When plants that are vernalization-requiring (*FRI*+) were grown without vernalization, *swn-/-* plants flowered later than *SWN*+ plants. The data also suggests a dosage dependency as *FRI*+/+ *swn*-3/-3 plants flower later than *FRI*+ *SWN*+ and *FRI+/+swn-3/+* plants showed an intermediate phenotype. This suggests that *SWN* regulates flowering independently of vernalization acting *FLC* and/or *FRI* though the autonomous pathway.

These novel discoveries about *SWN* further extend the developmental importance of the Pc-G complex in *Arabidopsis* and further discussion and expression analysis continues in chapter 4.

The following chapter pays attention to understanding the *swn- clf*- phenotype.

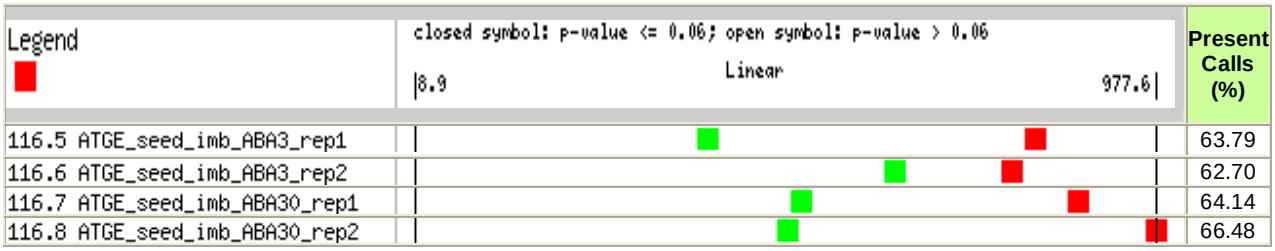

Fig. 2. 1. The effect of ABA on *SWN* and *CLF* expression.
Microarray data from in wild-type seedlings after 3 minutes (116.5 ATGE_seed imb_ABA3_rep1 and 116.6 ATGE_seed imb_ABA3_rep2) and 30 minutes (116.7 ATGE_seed imb_ABA30_rep1 and 116.8 ATGE_seed imb_ABA30_rep2 minutes) of ABA exposure. It shows that *SWN* (At4g02020) expression is relatively higher than *CLF* (AT2g23380). Closed symbols indicate present calls (P = < 0.06; "signals significantly higher than background") on a linear scale. Red squares = *SWN*. Green squares = *CLF*. (Figure is taken from the "digital northern" tool on the https://www.genevestigator.ethz.ch/at/ (Zimmermann *et al.*, 2004)).

Treatment: NaCl (150mM) 15 days old				
		Root length (mm)	**Hypocotyl length (mm)**	**% Germination**
Genotype	**wild-type**	14.5 ± 0.76	1.9 +/- 0.6	100
	swn-3	15 ± 0.79	2.1 +/- 0.7	100

Table. 2. 1. The effect of salt stress on germination and seedling growth in *swn-3* mutants.
Root and hypocotyl length and percentage germination were recorded every three days. No difference was observed between wild-type and *swn-3* mutant germination, root or hypocotyl growth. Wild-type is Columbia-0 (Col-0), *swn-2*, *swn-3*, and *swn-7* are in *Col-0* background. (n = 28 for the root and hypocotyl measurements, germination: wild-type n = 44, *swn-3* n = 37). Standard error is indicated and are used throughout)

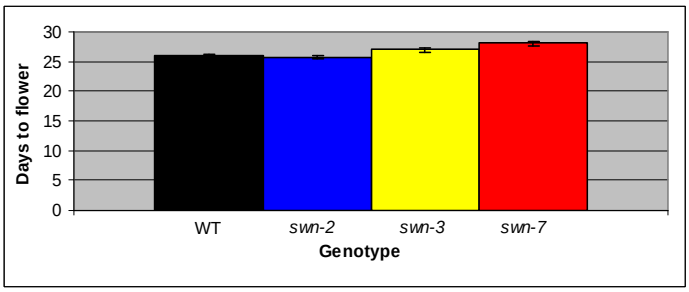

Fig. 2. 2. The effect of *swn*- mutations on flowering time in long days.
Flowering time was measured in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines by the number of days it taken to form a 1 cm bolt. No difference in flowering time was found between wild-type, *swn-2*, *swn-3*, or *swn-7* in the number of days to flower. (n = 20 for each genotype) (error bars indicate the standard error of means in this and subsequent figures).

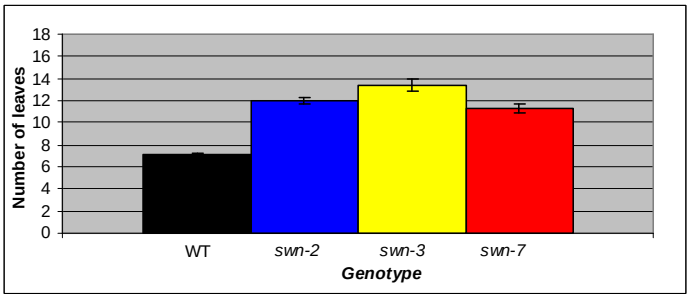

Fig. 2. 3. The effect of *swn*- mutations on leaf number at the time of flowering in long days.
The number of leaves at the time of a 1cm bolts in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines. Students t-test shows that *swn-2*, *-3*, and *-7* have significantly more leaves than wild-type (P <0.001) (n = 20 for each genotype).

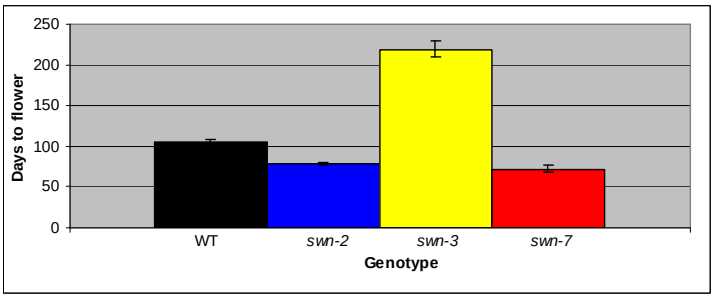

Fig. 2. 4. The effect of *swn*- mutations on flowering in short days.
The number of days it took to form a 1 cm bolt in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines. Wild-type plants flowered later than *swn-2* and *swn*-7 but *swn*-3 flowered later than wild-type. This shows high variability amongst *swn*- genotypes with respect to flowering time. (n = 20 for each genotype).

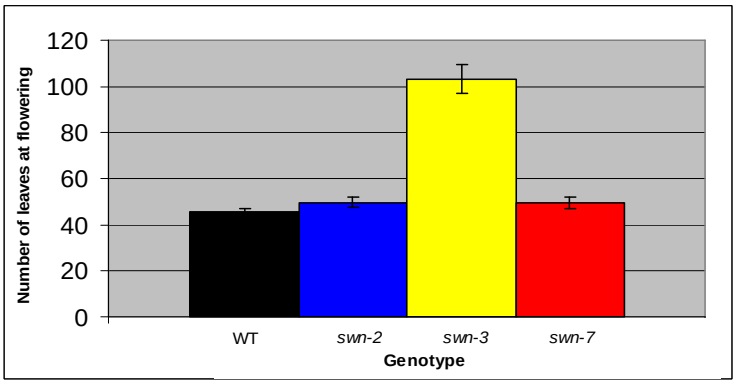

Fig. 2. 5. The effect of *swn*- mutations on number of leaves at flowering in short days.

The number of leaves was scored prior to flowering (1 cm bolt) in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines. A Students t-test showed that *swn-7* plants flowered with statistically more leaves than wild-type ($P < 0.06$), and *swn-2* and *swn-3* flowered with more leaves than wild-type ($P < 0.05$) (n = 20 for each genotype).

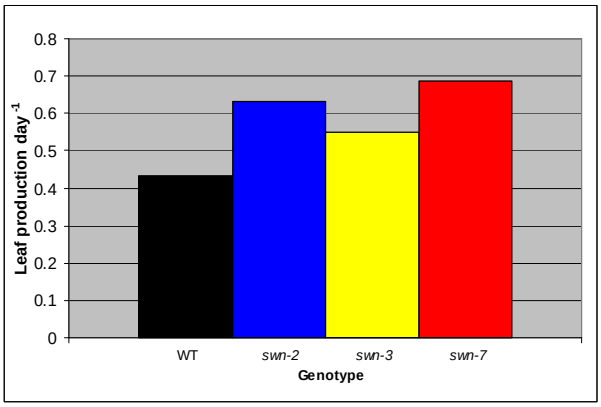

Fig. 2. 6. The effect of *swn*- mutations on leaf initiation rate in short days.

Leaf initiation rate was calculated by dividing the number of leaves at flowering and the number of days it took to flower, resulting in number of leaves produced per day. *swn-2*, *swn-3* and *swn-7* showed increased leaf production rate compared to wild-type (WT) (n = 20 for each genotype)

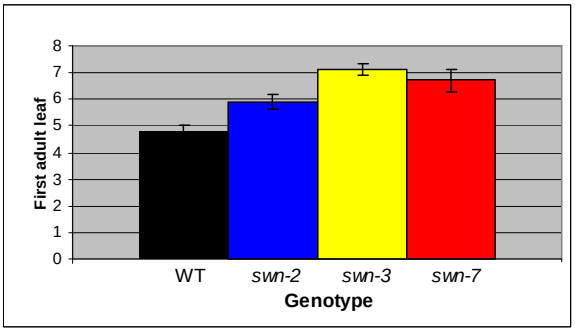

Fig. 2. 7. The effect of *swn*- mutations on juvenile leaf number in long days.

The number of juvenile leaves (as judged by the absence of abaxial trichomes) in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines were compared. A Students t-test showed that each *swn*- line, on average, possessed significantly more juvenile leaves compared to wild-type (P <0.01) (n = 15 for each genotype).

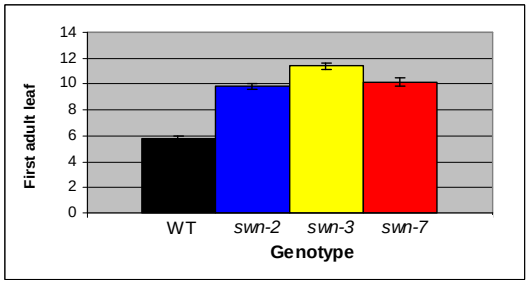

Fig. 2. 8. The effect of *swn*- mutations on juvenile leaf number in short days.
The number of juvenile leaves (as judged by the absence of abaxial trichomes) in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines were compared. A Students t-test showed that each *swn*- line, on average, possessed significantly more juvenile leaves compared to wild-type (P <0.01) (n = 20 for each genotype).

Day length	Genotype	Leaf initiation day $^{-1}$	Number of juvenile leaves	Juvenile to adult transition (days)	Difference between WT and *swn*- lines (days)
Long	**WT**	0.66	4.82	**7.3**	**0**
	swn-2	0.67	5.90	**8.9**	**+ 1.6**
	swn-3	0.71	7.10	**9.9**	**+ 2.6**
	swn-7	0.84	6.71	**8.0**	**+0.7**
Short	**WT**	0.32	5.81	**18.4**	**0**
	swn-2	0.37	9.85	**26.5**	**+ 8.1**
	swn-3	0.42	11.40	**27.1**	**+ 8.7**
	swn-7	0.39	10.15	**26.1**	**+ 7.7**

Table. 2. 2. The effect of *swn*- mutations on the juvenile to adult transition.
Leaf initiation rates were taken from the juvenile phase of development in wild-type (WT), *swn-2*, *swn-3* and *swn-7* lines in long and short days. The time it took to undergo the J-A transition in days was calculated by dividing leaf initiation rate by number of juvenile leaves. In both long and short day conditions, each *swn*- line underwent the J-A transition later than wild-type. In long days *swn*- lines underwent the J-A transition a minimum of 0.7 days later than wild-type. In short days *swn*- lines underwent the J-A transition a minimum of 7.7 days later than wild-type. This shows that the J-A transition delay is enhanced under short day conditions. (n = 20 for each genotype).

Genotype	Wild-type	swn-3/+	swn-3/-3
Expected: is *swn-3/-3* is not responsible	6.25	12.5	6.25
Expected: if *swn*-3/-3 is responsible	0	0	25
observed	4	8	13

Table. 2. 3. The effect of *swn-3* on J-A transition in long days.
A population of 78 plants segregating *swn-3* mutation were scored and 25 plants genotyped as they were judged to have a "mutant" phenotype (≥10 juvenile leaves). If *swn-3 /-3* were fully responsible for the increased number of juvenile leaves, then 100% of plants found with the highest number of juvenile leaves would be *swn-3/-3*. If the *swn-3/-3* genotype were independent of the increased number of juvenile leaves then a quarter of the plants genotyped would be *swn-3/-3* due to segregation. To test analyze whether the *swn-3/-3* mutants occurred more frequently than expected a χ^2 test was carried out. The χ^2 test showed that *swn-3/-3* plants were found to have significantly more than expected if *swn-3/-3* had no role in regulating the number of juvenile leaves ($P < 0.01$).

Genotype	wild-type	swn-3/+	swn-3/-3
Expected: is *swn-3/-3* is not responsible	3.75	7.5	3.75
Expected: if *swn*-3/-3 is responsible	0	0	15
Observed	2	4	9

Table. 2. 4. Effect of *swn-3* on the J-A transition in short days.
A population of 92 plants segregating *swn-3* mutation were scored and 15 plants as they were judged to have a "mutant" phenotype (≥22 juvenile leaves) and were genotyped. If *swn-3 /-3* were fully responsible for the increased number of juvenile leaves, then 100% of plants found with the highest number of juvenile leaves would be *swn-3/-3* plants. If the *swn-3/-3* genotype were independent of the increased number of juvenile leaves then a quarter of the plants genotyped would be *swn-3/-3* due to independent segregation. To test whether the *swn-3/-3* mutants occurred more frequently than expected a χ^2 test was carried out. The χ^2 test showed that *swn-3/-3* plants were found significantly more than expected if *swn-3/-3* had no role in regulating the number of juvenile leaves ($P < 0.01$).

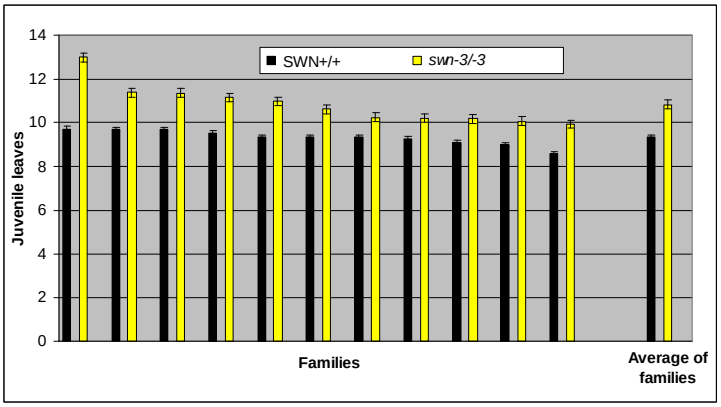

Fig. 2. 9. The effect of *swn-3/-3* on number of juvenile leaves in long days.
The averages of eleven families of *SWN+* and *swn-3/-3* were scored for number of juvenile leaves in long days. Each of the eleven *swn-3/-3* families possessed more juvenile leaves than the eleven *SWN+* families. A student t-test showed the average of the eleven families of *swn-3/-3* had significantly more juvenile leaves than the average of the eleven *SWN+* families ($P < 0.001$). (n = 11 families, 12 plants family^{-1})

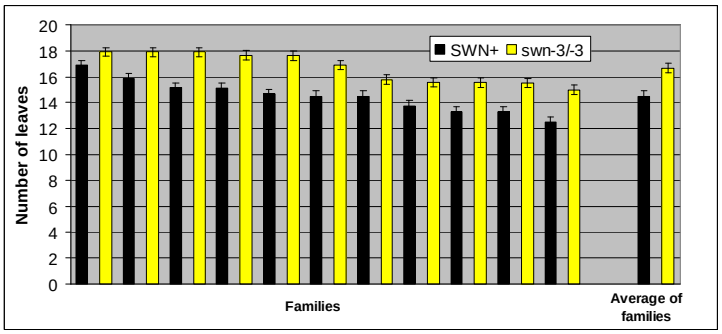

Fig. 2. 10. The effect of *swn-3/-3* on number of leaves at flowering in long days.
The average number of leaves at flowering in eleven F3 *swn-3/-3* and *SWN+* families grown in long day conditions. The data is ordered in descending order for each genotype. 6/11 families of *swn-3/-3* were found to have more leaves than *SWN+* lines. A student t-test showed that the average leaves from the eleven *swn-3* families is significantly more than the *SWN+* families ($P < 0.001$). (n = 11 families, 12 plants family^{-1}).

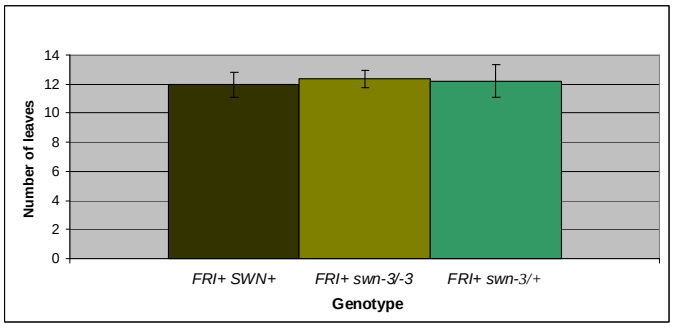

Fig. 2. 11. The effect *swn-3* on the vernalization response in long days.
The average number of leaves at flowering was scored in lines exposed to 6 weeks cold treatment (4 ℃) and grown in long days. *FRI+/+* plants flower with >100 leaves without vernalization treatment, showing the vernalization treatment was effective. No difference was observed between the average number of leaves at flowering in *FRI+SWN+* and *FRI+ swn-3/swn-3* plants. (n = 30 for each genotype).

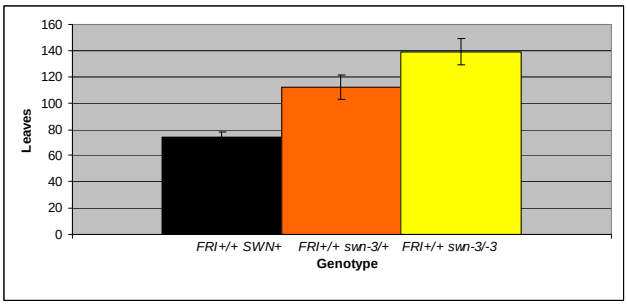

Fig. 2. 12. The effect of *swn-3* on flowering time in a *FRI+* background without vernalization.
The average number of leaves of F3 families homozygous for *FRI+/+* and segregating *swn-3* (*FRI+/+ swn-3/+*) were grown and genotyped. This shows a dosage dependant effect of *SWN* on the number of leaves produced prior to flowering, with *FRI+/+ SWN+* plants flower with fewer leaves than *FRI+/+ swn-3/+* or *FRI+/+ swn-3/-3*. Student t-tests showed that *FRI+/+ SWN+* has significantly less leaves than *FRI+/+ swn-3/+* or *FRI+/+ swn-3/-3* ($P < 0.01$). Student t-test showed that *FRI+/+* swn-3/+ had significantly less leaves than *FRI+/+ swn-3/-3* ($P < 0.01$). (n = 15 each genotype).

3. 0. Morphological analysis of *swn- clf-* and *vrn2- emf2-* mutants

swn- and *clf-* mutants show no severe disruption of root growth, and both *SWN* and *CLF* are found strongly expressed in the root (Wang *et al.*, 2006; Chanvivattana *et al.*, 2004; Schubert *et al.*, 2006). The phenotypes of loss of function *swn- clf-* and *vrn2-1 emf2-3* double mutants are similar, showing severely disrupted root and shoot development. The root exhibit a "pickle" root phenotype (swollen, opaque, and stunted growth). No true aerial organs are formed, and cell proliferation occurs in the meristematic regions eventually giving rise to callus-like tissue following germination. After several weeks or months growth, organ-like structures are formed on the surface of this callus as well as the occasional somatic embryo (Fig. 1. 6) (Chanvivattana *et al.*, 2004). *swn-1* is a phenotypically weak mutant allele and produces full-length transcript but at largely reduced levels (personal communication Dr. J. Goodrich). The *swn*-1 *clf-50* double mutant exhibits a weak phenotypic intermediate between the strong double mutants (*swn-7 clf-81* or *swn-7 clf-28*) and *clf-* plants. *swn-1 clf-50* mutants are smaller, with tightly curled leaves, very early flowering, and the floral defects are enhanced, compared to *clf-* mutants (Chanvivattana *et al.*, 2004). (Fig. 1. 6). No root transformation occurs, although the roots are shorter than wild-type roots (data not shown).

The *swn- clf-* phenotype raised several questions, which are tackled in this chapter: First, which processes cause the *swn- clf-* root deformation? I aimed to test if the loss of *SWN* and *CLF* leads to the loss of root identity, resulting in the callus-like tissue.

Secondly, do *SWN* and *CLF* function during embryogenesis? I was also interested in whether defects are present during embryogenesis in *swn- clf-* plants, as *SWN* and *CLF* are expressed during embryogenesis. Moreover, defects appear soon after germination in organs formed during embryogenesis. *SWN* is known to be involved in seed development where it acts redundantly with *MEA* to repress endosperm proliferation, as *swn-* in a *mea-* background enhances the *mea-* autonomous endosperm (Wang *et al.*, 2006). Whereas, *CLF* appears not to act redundantly with *MEA*, as the *clf- mea-* phenotype is identical to the *mea-* phenotype with respect to autonomous endosperm

(personal communication Dr. U. Grossniklaus). Therefore, I investigated whether *SWN* and *CLF* act redundantly during embryogenesis, particularly in the embryo, as the FIS2-Pc-G is only present in the endosperm (Kohler *et al.*, 2003a) suggesting a role for the EMF2-Pc-G to act during embryogenesis in the embryo.

3. 1. Do *swn- clf-* embryos show any morphological or expression defects during embryogenesis?

To investigate whether *SWN* and *CLF* act redundantly during embryogenesis, *swn- clf-* embryos were analyzed for phenotypic differences compared to wild-type. This was achieved by examining developing embryos in cleared seeds of *swn-7/-7 clf-28/+* plants, segregating 1:4 *swn-7/-7 clf-28/-28* double mutant embryos. No morphological differences were observed in any embryos or seeds in segregating siliques (Fig. 3. 1).

Although *swn- clf-* embryos have wild-type morphology it does not rule out that direct targets of the Pc-G complex are ectopically expressed in *swn- clf-* embryos. In wild-type plants, *AP1* is specifically expressed in flowers and is required for the identity of sepals and petals (Jack *et al.*, 1994). The promoter fusion of *AP1* to β-glucuronidase (*GUS*) reporter (*AP1::GUS*) is also found specifically in flowers. When the *AP1::GUS* is expressed in *clf-* and *emf2-* mutants *AP1::GUS* can be found ectopically expression in after germination in the leaves (Chanvivattana *et al.*, 2004). *swn- clf-* also show ectopic *AP1::GUS* in seedlings, specifically in the hypocotyl (personal communication Dr. O. Clarenz (Fig. 3. 2)). The embryos produced by self pollinating *swn-7/-7 clf-28/+ AP1::GUS* plants were analyzed to determine if ectopic *AP1* expression occurs during embryogenesis. No ectopic *AP1::GUS* expression was detected in any of the embryos (Fig. 3. 2). This suggests that both morphology and gene expression in *swn-7/-7 clf-28/-28* seeds is normal.

Although *SWN* and *CLF* are both expressed in embryogenesis (Wang *et al.*, 2006; Chanvivattana *et al.*, 2004; Goodrich *et al.*, 1997), they appear not to be required during embryogenesis, although they are crucial following germination. This may suggest that

SWN CLF have different roles in embryogenesis, or that *MEA* is acting to mask the effect of lacking *SWN* and *CLF* in embryogenesis. This may imply that *SWN, CLF,* and *MEA* act redundantly during embryogenesis. The true role of *SWN, CLF* in embryogenesis may only be revealed in the *swn- clf- mea-* triple mutant, however, *mea-* embryos are early embryo lethal, which would hamper this analysis.

As the morphology of the *swn- clf-* embryos appear normal, I further characterized the dramatic post-germinative phenotype of plants lacking Pc-G function.

3. 2. Analysis of external morphology of *swn- clf-* and *vrn2- emf2-* seedlings using scanning electron microscopy

To investigate the morphology of the root tip and epidermis of *swn-7 clf-28,* and *vrn2-1 emf2-3* mutants to wild-type seedlings scanning electron micrographs (SEM) were taken at 4 days and 14 days old.

At four days, *swn-7 clf-28* and *vrn2-1 emf2-3* seedlings were smaller with blunted root tips (Fig. 3. 3 F) compared to wild-type (Fig. 3. 3 B). The root tip phenotype was most prominent in the transformed root tips of *swn-7 clf-28* mutants (Fig. 3. 3 I). Root hairs are normally initiated in the differentiation zone above the cell division and expansion zone and orientated in alternate cell files of in the trichoblast cells, and are absent from atrichoblast cells (Dolan *et al.*, 1993). *swn-7 clf-28* mutants showed root hairs in contiguous cell files (Fig. 3. 3 Q), and were often deformed, being bulbous or swollen at the proximal end and normal at the distal; others were bulbous from base to tip (Fig. 3. 3 G, H, and P). The root hair defects mostly occurred at the root-hypocotyl junction (Fig. 3. 3 F). *swn-7 clf-28* seedlings also showed infrequent indeterminate out-growths on the otherwise normal cotyledon surface (Fig. 3. 3 E, M, and L). None of the defects described were observed in wild-type plants (Fig. 3. 3 A-C).

After 14 days growth *swn-7 clf-28* and *vrn2-1 emf2-3* seedlings showed the same blunted root tips and reduced cell size (Fig. 3. 3 I). The root hair positioning and

morphology defects were more common, and increased in severity after 14 days. In addition, root hair initiation defects were observed as well as the root hair positioning and morphology defects (Fig. 3. 3 I). *swn-7 clf-28* and *vrn2-1 emf2-3* mutants show root hair initiation above already established root hairs outside of the differentiation zone (Fig. 3. 3 Q), which are not found in wild-type plants (Schiefelbein & Somerville, 1990)(schematic depicting phenotype Fig. 5. 2).

The aerial parts of *swn-7 clf-28* and *vrn2-1 emf2-3* also showed increased deformities after 14 days. A plethora of new mal-formed lateral structures, are found in *swn-7 clf-28* and *vrn2-1 emf2-3* apical meristem, they are possibly leaf-like judging by morphology and whorled patterning (Fig. 3. 3 J + K). The leaf-like structures are deeply serrated, and form irregular projections on the edges and blade (Fig. 3. 3 K + M). The multi-cellular projections can be either spherical or linear in shape, almost root-like (Fig. 3. 3 E, K, L, and R). Other projections were uni-cellular and spherical in shape (Fig. 3. 3 E, K, L, and R). Irregular cellular proliferation also occurred at the hypocotyl-root junction, where large rounded cells are found (Fig. 3. 3 G, H, and P) (schematic depicting phenotype Fig. 5. 2).

Wild-type epidermis of leaves and cotyledons consists of inter-digitated cells reminiscent of a jigsaw pattern (Yang & Sack, 1995). The lateral organs of *swn-7 clf-28* and *vrn2-1 emf2-3* lacked this jigsaw patterning (Fig. 3. 3 E, L, M, and R). Irregularly sized and shaped stomata were also observed. Stomata were incredibly large, approximately three times the size of wild-type stoma, and others were prominent from the surface (Fig. 3. 3 R). Stomata positioning was only rarely incorrect in *swn-7 clf-28* and *vrn2-1 emf2-3* (Fig. 3. 3 E + R), the genetically controlled one cell spacing rule between stomata guard cells was frequently observed (Yang & Sack, 1995). These defects were not observed in wild-type seedlings (Fig. 3. 3 C).

Interestingly, the cotyledons and root epidermides of *swn-7 clf-28* and *vrn2-1 emf2-3* seedlings showed adhesion defects, or holes, at cell junctions which was not observed in

wild-type plants (Fig. 3. 3 M, N, and O). This indicates that *SWN* and *CLF* play a role in epidermis cell wall regulation.

SEM analysis of *swn-7 clf-28* and *vrn2-1 emf2-3* mutants unveiled previously unknown developmental defects, which were identical, indicating potential novel functions of the Pc-G complex. These include regulating epidermis differentiation, maintenance of meristems, regulating atrichoblast differentiation, root hair initiation and development, stoma development, maintaining cell cohesion and regulating cell expansion (schematic depicting phenotype Fig. 5. 2).

The growth projections found in *swn-7 clf-28* and *vrn2-1 emf2-3* seedlings were ill defined, suggesting the beginning of callus-like tissue. This suggests that cells in the cotyledon and newly formed cells have changed their differentiation state. This subsequently allows aberrant cell division in these tissues, which may go on to form the callus-like tissue. *swn-7 clf-28* and *vrn2-1 emf2-3* plants analyzed did not produce any somatic embryo, which has been previously observed in "mature" *swn-3 clf-50* callus-like tissue (Chanvivattana *et al.*, 2004). This suggests that the *swn-7 clf-28* and *vrn2-1 emf2-3* phenotypes depend on developmental age to show the full range of defects. This suggests that somatic embryo formation is likely a down-stream consequence, rather than direct cause of loss of *SWN* and *CLF* function. However, the callus tissue of *swn-7 clf-28* or *vrn2-1 emf2-3* was not analyzed (schematic depicting phenotype Fig. 5. 2).

3. 3. Analysis of internal root tip morphology in *swn- clf-* and *vrn2-emf2-* mutants

The root tip is composed of distinct cell types, including: epidermis, endodermis, cortex, pericycle, vasculature, quiescent centre, stem cells and the collumella cells (Fig. 3. 4 A). Each of these cell types has distinct size, shape, orientation, and gene expression (Dolan *et al.*, 1993; Brady *et al.*, 2007). It is generally accepted that each cell type is formed from the actively dividing stem cell population in the root tip (Dolan *et al.*, 1993). The stem cell population is maintained in the undifferentiated state by signalling from the Quiescent Centre (QC) cells (Dolan *et al.*, 1993). The QC cells are four non-dividing

cells in the root tip surrounded on every side by the stem cell pool (Dolan *et al.*, 1993). When a stem cell divides, one daughter cell is maintained as a stem cell and the other daughter cell differentiates into one of the root cell files (Dolan *et al.*, 1993). The differentiation of stem cells involves cell type specific expression, protein movement and the complicated interaction of proteins in different complexes resulting in the specification of differentiation states. (van den Berg *et al.*, 1995; van den Berg *et al.*, 1997; Dolan *et al.*, 1993). For example, the transcription profile of endodermis cells is determined by a GRAS transcription factor, *SCARECROW* (*SCR*). *scr*- mutants display disrupted radial root patterning, lack cell layers, and the resultant cell file was of a mixed cortex/endoderm fate (Di Laurenzio *et al.*, 1996; Sabatini *et al.*, 2003). *scr*- mutants cannot determine or maintain the endodermis during embryogenesis, or if knocked out post-germination (Heidstra *et al.*, 2004). *SCR* is also expressed in the QC, and the reduced *SCR* expression causes disruption of the QC, stem cell population, and the asymmetric cell division which give rise to the cell files (Di Laurenzio *et al.*, 1996; Heidstra *et al.*, 2004; Sabatini *et al.*, 2003). *SHORT ROOT* (*SHR*) determines the cortex cell layer, as demonstrated by the *shr*- mutants lacking the cortex cell layer (Helariutta *et al.*, 2000). *SHR* is expressed in the cortex but it is capable of moving in to the vasculature, and the endodermis to probably positively regulate *SCR* expression (Helariutta *et al.*, 2000; Cui *et al.*, 2007; Nakajima *et al.*, 2001a). *JACKDAW* and *MAGPIE* zinc finger proteins can interact in multimeric complexes with *SCR* and *SHR* to determine the fate of the distinctive cell files and asymmetric cell division (Welch *et al.*, 2007). However, the mechanism of regulating tissue specific transcriptional profiles is still largely unknown.

Given the severe root deformities, it raised the question whether they were a consequence of disrupted QC function, and this was first explored by using confocal microscopy to observe if the root architecture was disrupted in Pc-G mutants. To investigate the internal cellular morphology of the deformed *swn- clf-* roots they were stained with Prodidium Iodine (PI) and confocal microscopy was carried out on different developmental stages. PI stains cell walls and dead cells. Figure 3. 4 reveals that root

tips of 14-day-old *swn-7 clf-81* seedlings had blunted root tips and the cells surrounding the QC were smaller, irregular in shape and orientation (Fig. 3. 4 E) compared to wild-type roots (Fig. 3. 4 B). The same features were found in the roots of *vrn2-1 emf2-3* seedlings (Fig. 3. 4 D). The weaker *swn-1 clf-50* double mutant also showed the same features but to a lesser degree (Fig. 3. 4 C).

Interestingly, each of the mutant lines analyzed showed morphologically normal cell files above the QC, but the root cap possessed increased numbers of cells (Fig. 3. 4).

Four day old mutants of *swn-7 clf-81*, *vrn2-1 emf2-3*, and *swn-1 clf-50* also showed blunted root tips and cell size and orientation defects, to a lesser extent than at 14 days old. The cell files were ordered superior to the root tip. Similarly, *swn-7 clf-28* mutants with transformed root tips show cell size and orientation defects in the root tip, but had normal cell files after 4 days. The root phenotypes of each mutant line showed variation, but were replicable. No gross phenotypic differences were observed in either *swn-* or *clf-* roots compared to wild-type roots (data not shown).

These results suggest that the severity of the root phenotype increases with time after germination. This is certainly the case for formation of the callus-like tissue, which increases over time, strongly suggesting the *swn- clf-* root deformity is a consequence of the newly formed cells rather than the cells formed from embryogenesis, or the loss of cell identity takes time to occur which is reflected in the phenotype.

The evidence that the internal morphology in plants lacking Pc-G function is disrupted led me to investigate whether the morphological defects were a consequence of differentiation defects in the stem cell niche.

3. 4. Analysis of *SHR* and *SCR* expression in *swn- clf-* and *vrn2-emf2-* mutants

The QC and stem cells were of particular interest because these cell types showed defects in the root tips of *swn- clf-* and *vrn2-1 emf2-3* mutants. *SCR* and *SHR* define endodermis and QC identity, by their specific expression patterns and protein locality

(Heidstra *et al.*, 2004; Helariutta *et al.*, 2000; Nakajima *et al.*, 2001b). I aimed to test whether the Pc-G complex controls internal root development by regulating *SCR* and/or *SHR* expression; this was achieved by analysing *SCR::GFP* (Green Fluorescent Protein) and *SHR::GFP* promoter fusions lines in *swn-7 clf-81, swn-1 clf-50* and *vrn2-1 emf2-3* backgrounds. *SCR* possesses the H3K27me^3 mark and is therefore a likely target of the Pc-G complex. If *SWN CLF* is directly responsible for *SCR* and *SHR* expression, then ectopic expression of *SCR* and *SHR* might be found in *swn- clf-* plants, potentially causing the *swn- clf-*phenotype.

Figure. 3. 5 shows that after 4 days there were no distinguishable differences between wild-type and *swn-7 clf-81, swn-1 clf-50, vrn2-1 emf2-3* plants in expression of *SCR::GFP* or *SHR::GFP*. This finding is in parallel with the cell files, cell shapes, and sizes being relatively normal after four days in *swn-7 clf-81, swn-1 clf-50,* and *vrn2-1 emf2-3* lines. *swn-7, swn-1, clf-50, clf-81, emf2-3* and *vrn2-1* single mutants also showed wild-type expression of *SCR::GFP* and *SHR::GFP* (data not shown).

After 14 days *swn-7 clf-81* and *vrn2-1 emf2-3* roots showed expression of *SCR::YFP* (Fig. 3. 5 C + E) and *SHR::GFP* (Fig. 3. 5 H + J) was lost or massively reduced compared to wild-type seedlings. The *swn-1 clf-50* mutant showed a slight reduction of expression of *SCR::GFP* (Fig. 3. 5 B + D) and *SHR::GFP* (Fig. 3. 5 G + I) compared to wild-type (Fig. 3. 5 A + F). However, *swn-7, swn-1, clf-50, clf-81, emf2-3* and *vrn2-1* single mutants showed no difference to wild-type expression of *SCR::GFP* and *SHR::GFP* after 14 days (data not shown).

Despite the variation in intensity of expression, at 4 and 14 days, the cell specific expression pattern of *SCR::YFP* and *SHR::GFP* in *swn-7 clf-81, swn-1 clf-50, vrn2-1 emf2-3* was identical to wild-type plants. This shows us that the *swn- clf-* and *vrn2- emf2-* phenotype is not a likely consequence of ectopic or over-expression *SHR* or *SCR*. The changes in *SHR* and *SCR* expression are likely secondary effects through mis-expression of other genes regulating *SHR* and *SCR* expression. *scr- shr-* plants have reduced growth, and root morphology defects found in embryos that lack endodermis

and cortex cells (Helariutta *et al.*, 2000), probably as a consequence of losing the QC identity. The reduced growth is comparable to the effects observed in *swn- clf-* and *vrn2- emf2-* seedlings, but by no means identical. The decreased expression of these important transcription factors could explain the reduced growth. However, the other aspects of the *swn- clf-* and *vrn2- emf2-* phenotypes are probably not due to *SHR* and *SCR* down regulation.

Auxin distribution is another method employed to maintain the stem cell population and considering: A) the pleiotropic function of auxin in plant development and the pleiotropic phenotype of *swn- clf-* mutant (Chanvivattana *et al.*, 2004; Kim *et al.*, 2007). (B) The similarity of the *swn- clf-* is reminiscent of wild-type plants grown on Callus Inducing Media (CIM) both causing callus material, it provides good reasoning that auxin distribution is a likely candidate to cause of the *swn- clf-* mutant root deformation.

3. 5. Analysis of auxin distribution and concentration in *swn- clf-* and *vrn2-emf2-* mutants

A further method was used to investigate the root phenotype: auxin distribution. Auxin plays a central role in forming the root in embryogenesis, and maintaining the root post-germination, by controlling gene expression to maintain the stem cell population, and regulate division and elongation (Teale *et al.*, 2006). An example of this point is demonstrated elegantly as auxin maxima in the root tip (Grieneisen *et al.*, 2007). The *DR5::GUS* and *DR5::GFP* expression is an indirect indicator of auxin concentration, it is a synthetic promoter comprised of five repeats of auxin responsive binding element found in *AUXIN RESPONSE FACTORS* fused to either *GUS* or *GFP* reporters (Ulmasov *et al.*, 1997; Friml *et al.*, 2003). Expression in wild-type seedlings can be seen in Figure. 3. 6 (A + C). Auxin distribution and concentration was analyzed in *swn-1 clf-50,* and *vrn2-1 emf2-3* lines by observing *DR5::GUS* and *DR5::GFP* expression, i.e. in plants with reduced (*swn-1 clf-50*) and lacking Pc-G function (*vrn2-1 emf2-3*). The *swn-7 clf-28* and *vrn2-1 emf2-1* phenotypes are virtually identical, suggesting that this is indeed representative of removing total Pc-G function post germination.

After 4 days growth *swn-1 clf-50,* and *vrn2-1 emf2-3* lines showed a relatively normal auxin distribution and concentration in root and aerial parts, (Fig. 3. 5 D + H), as did *swn-1, clf-50, emf2-3,* and *vrn2-1* single mutants (data not shown). After 14 days, *DR5::GUS* expression in *swn-1 clf-50* mutants is almost ubiquitous expression in the aerial parts (Fig. 3. 5 C + E), suggesting an increase in auxin concentration compared to wild-type plants. Conversely, there is a reduction in auxin concentration in the root tip (*DR5::GFP* expression) in *swn-1 clf-50,* and *vrn2-1 emf2-3* mutants (Fig. 3. 6 E + J). A patchy auxin distribution is seen throughout the *vrn2-1 emf2-3* mutants (Fig. 3. 6 G + I). No difference was observed in *swn-1, clf-50, emf2-3,* and *vrn2-1* single mutants compared to wild-type plants carrying the reporter constructs (data not shown).

The differential auxin concentration in the aerial parts and roots may reveal a possible reason for the difference in the phenotype of the above and below ground parts in early development in *swn- clf-* and *vrn2-1 emf2-3* mutants. Mutating genes responsible for auxin biosynthesis, i.e. *YUCCA2* (*yuc2-*) *yuc4-* causes failure to develop a root meristem in embryos (Cheng *et al.*, 2007). Auxin is also required to maintain the competence for root growth (Grieneisen *et al.*, 2007). Therefore decreased auxin concentration in *swn- clf-* and *vrn2-1 emf2-3* roots might explain the reduced root growth. However, it cannot explain the blunted root tip, root transformation, or the root hair morphology or positioning defects in *swn- clf-* and *vrn2-1 emf2-3* roots. The increased auxin concentration, as seen in the over-expression of *YUC2* and *YUC4* in the aerial parts show epinastic cotyledons, increased apical dominance, and curled leaves (Cheng *et al.*, 2007; Kim *et al.*, 2007). This is not observed in the strong *swn- clf-* mutants due to the severity of the phenotype. The abnormal auxin distribution in *swn- clf-* mutants correlates with aberrant cell division possibly directly causing the abnormal outgrowths found on *swn- clf-* and *vrn2-1 emf2-3* mutants, particularly in the aerial parts. Curled leaves found in the weak *swn-1 clf-50* mutants could be due to the increased auxin concentration but it could also be due to the mis-expression of *AG* as is the case in *clf-* plants (Goodrich *et al.*, 1997), or indeed a combinatorial effect, that may justify the enhanced phenotype.

The high concentration of auxin in aerial parts, and lack of auxin in the roots in *swn-1 clf-28* and the patchy distribution in *vrn2-1 emf2-3* illustrates that ectopic expression is occurring in a tissue specific manner. This also implies the Pc-G is likely to play a role in the regulation of auxin in plant development. The differential auxin concentrations in the Pc-G mutants could be caused by the tissue specific mis-expression of targets of direct Pc-G targets i.e. secondary, indirect targets, or down stream genetic effects, rather than the direct Pc-G targets themselves.

These results illustrate the root deformation could be in part a consequence of mis-regulation of transcription factors and hormones. As the root phenotype of the *swn- clf-* and *vrn2-1 emf2-3* cannot be explained exclusively by the down regulation of *SCR, SHR,* and auxin concentration, it is therefore more likely other, possibly not root specific factors cause the root phenotype.

In essence, the reduced auxin concentration and reduction of *SHR* and *SCR* expression in *swn- clf-* and *vrn2-1 emf2-3* plants suggest that the tissues comprising the root have lost their root identity. In addition, the transformed regions of the root gain green colouration, suggesting acquisition of aerial parts identity. Therefore, it can be suggested the tissue of the root had lost their root identity. In which case, have the root tissues acquired a different differentiation state? Moreover, are there differences in the cell fates of Pc-G mutants and wild-type in the root cell files?

3. 6. Analysis of root differentiation in *swn-7 clf-28* seedlings and callus-like tissue

3. 6. 1. Have swn- clf- roots lost their stem cell niche?

To validate the hypothesis that the root identity of *swn- clf-* roots had changed starch, staining, using Lugol's stain, was carried out. Lugol's staining shows the presence of starch rich amyloplasts and statoliths. Statoliths are starch organelles in the root tip used for gravity response (Kiss *et al.*, 1989). The presence of statoliths in collumella cells

demonstrates a differentiated state. Undifferentiated stem cells and QC cells lack statoliths.

After 4 days the *swn-7 clf-81, vrn2-1 emf2-3* and *swn-1 clf-50* each showed an increase in the number of cells possessing starch bodies, extending into the stem cell population in the root tip compared to wild-type roots (Fig. 3. 7). This indicates that the root meristem identity is lost in these plants. In support of this, RNA *in situ* hybridization using QC25 (a QC specific probe) revealed that *swn-2 clf-81* roots showed less accumulation in the QC compared to wild-type in four-day-old seedlings (personal Communication, B. Scheres). It is possible that the cause of losing stem cell niche indeterminacy is a direct consequence of losing Pc-G function. This may explain the considerable reduction in *swn- clf-* and *vrn2-1 emf2-3* root growth. It may also explain why the *swn- clf-* and *vrn2-1 emf2-3* phenotype becomes strongly apparent after 4 days of growth, following the loss of stem cells, where the root tip cells divide much slower than wild-type, and in a much less organized pattern (Fig. 3. 4 D) resulting in the root tip abnormality.

When considering the increase in starch staining, and the reduced expression of *SCR::GFP*, *SHR::GFP*, *DR5::GUS* and *DR5::GFP* in *swn- clf-* and *vrn2-1 emf2-3* roots there appears to be a paradox. The root is both gaining and losing differentiation state. Reducing expression of *SCR* and *SHR* and reduced auxin concentration results in a loss of differentiation. Simultaneously however, there is an increase in differentiation as shown by starch granules being found in the stem cell population. However, this paradox is resolved when we realize that it is the stem cells which are gaining a differentiation state and the ground cells are losing/changing their differentiation state, so the cells gaining and losing differentiation state are occurring in different tissue types. I suggest that the Pc-G functions to maintain the identity of both stem cells and the differentiation states of other tissue of the root.

SCR expression is needed to maintain the QC identity, and the QC is required to maintain the stem cell niche in its undifferentiated state. Decreased *SCR* expression along with the decrease in auxin maxima could explain the loss of the stem cell population. The reduction in stem cell population is a likely consequence of QC losing its differentiation state. However, there is a time discrepancy between when the stem cell population is lost (after 4 days), and the decrease in *SCR* expression and auxin maxima is not observed until day 14. Therefore, it is unlikely that the loss of stem cells is a direct result of reduced *SCR* expression and/or auxin concentration alone. It is more likely other Pc-G regulated factors cause the loss of stem cells.

The Pc-G is fundamentally important for maintaining global root differentiation, partially through maintaining the QC and the stem cell population. The Pc-G probably does not regulate the QC identity through auxin distribution, or the *SHR-SCR* pathways. It suggests that the *swn- clf*-roots are losing their "root" identity, and so potentially acquiring a new differentiation state. If Pc-G function is lost then de-repressed genes from another differentiation state may be expressed to define its state. Which identity could the root tissues have taken on?

3. 6. 2. Have *swn- clf-* mutants acquired an embryonic cell fate?

swn-3 clf-50 callus material has been shown to produce somatic embryos (Chanvivattana *et al.*, 2004). This suggested that the *swn- clf-* root could be acquiring the differentiation state of embryogenesis. Other points of evidence that led to the hypothesis that the Pc-G acts to repress traits specific to embryogenesis in post-germination development: 1) The over-expression of a master regulator of seed maturation, *LEC1*, post-germination leads to the formation of somatic embryos on the cotyledons (Santos *et al.*, 2005). 2) *LEC1, LEC2* and *FUS3* were found up regulated in *swn- clf-* mutants (Makarevich *et al.*, 2006). 3) The "pickled" root phenotype of *swn- clf-* is virtually identical to *pkl-*, *PKL* is known to repress embryonic traits and genes mis-expressed in *pkl-* largely possess the histone methylation mark which is created by *SWN* and *CLF* (Zhang *et al.*, 2008).

A direct method to examine if a tissue has an embryonic fate is to stain tissues with Sudan 7. Sudan 7 stains red specifically for embryonic storage oils. *swn-7 clf-81, vrn2-1 emf2-3* and *swn-1 clf-50* at 4 and 14 days post germination, and *swn-7 clf-81* callus-like tissue were stained with Sudan 7. No staining was found in *swn-7 clf-81, vrn2-1 emf2-3* or *swn-1 clf-50* at four and fourteen days (Fig. 3. 8 E-G, I-K, M, N, and P-R) compared to wild-type seedlings (Fig. 3. 8 A-C). The *swn-7 clf-81* callus like tissue did show staining in sporadic patches independent of somatic embryos (Fig. 3. 8 D, H, L, and O). This suggests that the callus-like tissue has acquired the characteristics of the embryogenesis phase. This highlights that the "mature" *swn-7 clf-81* callus-like mutants do potentially acquire an expression profile similar to embryogenesis. The ectopic embryonic genes are unlikely to be a major factor contributing to the early *swn-7 clf-81, vrn2-1 emf2-3* or *swn-1 clf-50* root phenotype, as they lack embryonic storage oils in the seedling.

It maybe expected that the *swn-7 clf-28* callus-like tissue does accumulate embryonic storage oils, as it is known to produce somatic embryos. The complete embryonic expression gene networks may well be expressed. The early seedlings of *swn- clf-* and *vrn2-1 emf2-3* lines do not accumulate storage oils. The formation of storages oils is believed to be the output or at the bottom of the embryogenesis pathways (Wobus & Weber, 1999), thus it may take *swn- clf-* and *vrn2-1 emf2-3* mutant lines time to acquire the intricate transcriptional networks regulating embryonic characteristics. It illustrates a role for the Pc-G to repress the embryonic characteristics post-germination.

3. 6. 3. Is there a phenotypic effect of *swn- clf-* after growing on phyto-hormones?

It has been shown that the gene *PKL* is involved in repressing the embryonic specific expression, probably through chromatin remodelling (Li *et al.*, 2005; Dean Rider S Jr *et al.*, 2003). The *pkl-* root phenotype is reminiscent of strong *swn- clf-* root phenotype. The lateral root of *pkl-* seedlings transforms into a "pickle" root. It is an opaque blunted

root, which is stunted in growth (Fig. 3. 9). The “pickle” root does not develop any further. The aerial parts of *pkl-* plants are a darker shade of green, and are slightly later flowering in long and short days (Ogas *et al.*, 1999; Ogas *et al.*, 1997; Dean Rider S Jr *et al.*, 2003; Henderson *et al.*, 2004; Li *et al.*, 2005). The “pickle” root shows staining of storage oils after 4-10 days (Ogas *et al.*, 1999), whereas the *swn- clf-* mutants only show accumulation of storage oils after developing into callus-like material. The root transformation phenotypes of *pkl-* and *swn- clf-* are superficially identical (Fig. 3. 9), suggesting a similar genetic cause of the *pkl-* and *swn- clf-* root phenotypes. The root phenotype of *pkl-* mutants shows de-repressed, ectopic, expression of embryonic regulators i.e. *FUS3*, *LEC2* (Li *et al.*, 2005). These embryonic regulators control expression of other embryonic genes including those encoding enzymes that produce seed storage proteins and oils (Wobus & Weber, 1999). This explains the occurrence of storage oils found in the *pkl-* roots. This illustrates that *PKL* has a role in the repression of the embryonic regulators in post germination in the root tissues. There is only partial penetrance of the *pkl-* root transformation phenotype (~20%), the penetrance is increased to ~80% when *pkl-* seedlings are grown on media containing the GA inhibitor unicortazol (Henderson *et al.*, 2004). This illustrates that the *pkl-* “pickle” root phenotype is partially GA dependant, and that GA may be involved in repressing embryonic fate. When *swn-7 clf-81* or *swn-7 clf-28* mutants are grown on GA or a GA inhibitor (paclobutrazol) there was no difference in phenotype of the transformed root tip or indeed the penetrance of the transformed root tip was observed, which was observed for the *pkl-* mutants which was used as a control (data not shown).

As GA concentration has no influence on the *swn- clf-* phenotype it strongly suggests that the *swn- clf-* phenotype is a consequence of gene mis-regulation directly affecting cell fate. Thus, it is likely that the *swn- clf-* phenotype is independent of the *PKL* pathway, which is enhanced by reducing GA action but not necessarily dependant on GA activity. The Pc-G and CHD3 chromatin remodellers play a similar role in repressing embryonic differentiation post germination, probably by *SWN*, *CLF* and *PKL*

acting to repress key embryonic regulators. This point shall be explored further when comparing genome wide expression in each of the mutants in Chapter 4.

There are several other well documented pathways regulating gene expression in the seed maturation program including transcription factors and sucrose. Sucrose is known to play a role in promoting embryonic traits (Rook *et al.*, 2006). This lead to investigate whether sugar concentration could influence the *swn- clf-* phenotype.

3. 6. 4. *swn- clf-* mutants show phenotypic differences when grown on different sugar concentrations

Sucrose is a vital molecule in plants and documented to act in several different pathways in *Arabidopsis*, for example acting to promote master regulators of seed maturation such as *LEC2* (Tsukagoshi *et al.*, 2007). The *swn- clf-* double mutant phenotype is reminiscent of *HIGH SUGAR INDUCIBLE 2 (hsi2-) HIGH SUGAR EXPRESSION1 (hsl1-)* double mutants. *hsi2- hsl1-* show callus-like formation when grown on high sugar concentrations and accumulate embryonic storage oils (Tsukagoshi *et al.*, 2007). *HSI2 HSL2* genes encode B3 class transcription factors (Tsukagoshi *et al.*, 2007). When the *hsi2- hsl1-* double mutant is grown on media with a low sugar concentration (0%), they show swollen hypocotyls compared to wild-type (Tsukagoshi *et al.*, 2007). When *hsi2 hsl2* is grown on high sugar concentration (1%) roots over-proliferate and accumulate storage oils and proteins specific to the embryonic phase (similar to *pkl-),* and development is then halted and they die (Tsukagoshi *et al.*, 2007). The phenotypic similarity of *hsi2- hsl1-* and *swn- clf-* includes mis-regulation of cell division in the root, and accumulation of embryonic oils. The *hsi2- hsl1-* and *swn- clf-* phenotypes are far from identical as *hsi2- hsl1-* seedlings die after ~14 days (Tsukagoshi *et al.*, 2007) rather than being immortal like strong *swn- clf-* mutants (Chanvivattana *et al.*, 2004).

swn-2 clf-81 has a relatively strong phenotype but with a low penetration of root transformation. It was grown on a range of sugar concentrations and the penetrance of the transformed root tip phenotype was measured. After 14 days growth on 0%, 0.3%,

1% or 2% sucrose media *swn-2 clf-81* plants showed a transformed root tip penetrance of <5%. This increased to ~20% when grown on 4% sucrose (Fig. 3. 10).

It emerges that growing *swn-2 clf-81* plants on high sucrose concentration increases the severity and penetrance of its phenotype, which is similar to *hsi2- hsl1-* mutants. This further suggests that the root phenotype of strong *swn- clf-* mutants is correlated with increased embryonic traits.

3. 7. Summary and conclusions

Analysis of the *swn- clf-* double mutants has revealed a cacophony of developmental defects regulated by *SWN* and *CLF* acting redundantly post-germination. *swn- clf-* embryos are wild-type in morphology and most likely have normal gene expression of homeotic floral transcription factors like *AP1*. *MEA* is a known target of *SWN* and *CLF*, as *MEA* is found up regulated in *swn-3 clf-50* mutant seedlings (Jullien *et al.*, 2006) indicating that *MEA* may be up regulated in *swn- clf-* embryos. The presence of *MEA* in the embryo may mask the effect of the loss of *SWN* and *CLF*. This may imply that *SWN*, *CLF*, and *MEA* act redundantly during embryogenesis.

To evaluate the role of Pc-G in root development *swn-1 clf-50, swn-7 clf-81* and *vrn21-emf2-3* seedlings were analyzed using confocal and SEM microscopy. After 14 days growth, SEM revealed that *swn-7 clf-28* roots have blunted tips, and defects in root hair patterning and development, cell cohesion, cotyledons possessing multi-cellular out growths, and altered stomata development compared to wild-type plants. Overall suggesting the Pc-G has a role in maintaining a variety of cell fates post-germination. Confocal microscopy revealed that *swn- clf-* and *vrn2-1 emf2-3* mutant root meristems possess small, irregular shaped cells, abnormal cell division, and irregular cell orientation. The phenotype in each mutant line was enhanced with age.

The root stem cell population was found to be lost in plants lacking Pc-G function (*swn- clf-* or *vrn2-1 emf2-3* double mutants) after 4 days. Seedlings lacking Pc-G function also showed decreased expression of *SCR::GFP, SHR::GFP, DR5::GUS* and *DR5::GFP* after 14 days growth. This illustrates that the Pc-G is fundamental to the maintenance of the stem cell niche in the root, however, it is unlikely it does this through the direct regulation of auxin distribution, or the *SHR-SCR* pathway. The perturbed auxin distribution and decrease of SCR and SHR transcription factors could explain the retardation of root growth. It is unlikely that auxin, *SCR* or *SHR* are entirely responsible for the root phenotypes of the *swn-1 clf-50, swn-7 clf-81* and *vrn2-1 emf2-3* mutants, as reduced expression is likely to be a secondary effect, and their mis-regulation occurs after the deformation of the root tip, and loss of the stem cell undifferentiated state.

The down regulation of the root specific transcription factors *SCR* and *SHR*, and the loss of auxin in the root tip, in addition to the global phenotype, provide evidence the roots have lost their identity in *swn- clf-* and *emf2-3 vrn2-1* rather than singular aspects of root development. If this was the case then perhaps a different identity was being acquired.

Staining for embryonic oils revealed that four and 14 days old *swn-1 clf-50, swn-7 clf-81* and *vrn2-1 emf2-3* seedlings did not accumulate embryo specific oils in the roots. Interestingly, staining of *swn- clf-* "mature" callus-like tissue did show staining for embryonic oils. This shows that the Pc-G does have a role in the repressing embryonic specific expression post-germination, but *swn- clf-* mutants are slow to gain the full battery of the embryonic expression profile. In support of this other mutants display a similar phenotype to *swn- clf-* mutants have ectopic expression of embryonic traits.

The *swn- clf-* phenotype is reminiscent of *pkl-* phenotype. The penetrance of the *pkl-root* transformation is sensitive to GA concentration. When *swn- clf-* lines were grown on GA or GA inhibitor (paclobutrazol), no effect on phenotype or penetrance was observed. This suggests that the *swn- clf-* phenotype is not dependant on GA. The cause of the *pkl-* and *swn- clf-* phenotypes are therefore likely to be different. From this, we

can conclude that the Pc-G and the CHD3 chromatin remodellers function to repress embryonic expression post germination.

The *swn- clf-* phenotype has similarities to the *hsi2- hsl1-* mutant, which is involved in repressing embryonic traits after germination. *hsi2- hsl1-* double mutants show cell proliferation defects and accumulation of embryonic storage oils post germination on high sugar concentrations. The weaker *swn-2 clf-81* double mutant showed an increase in the penetrance of the transformed root tip when grown on high sugar concentration, suggesting the Pc-G function and sugar function interact. High sugar concentration is effecting transcription and this interacts with the mis-expressed genes in the *swn- clf-* seedlings, resulting in the enhanced transformation of the root tip.

This evidence leads to the conclusion that the plants lacking Pc-G function lose meristem identity, which may be correlated with ectopic expression of embryonic traits. A detailed analysis into genome transcription of the *swn-*, *clf-* and *swn- clf-* is evaluated in the following chapter permitting insight in the function of the Pc-G in development.

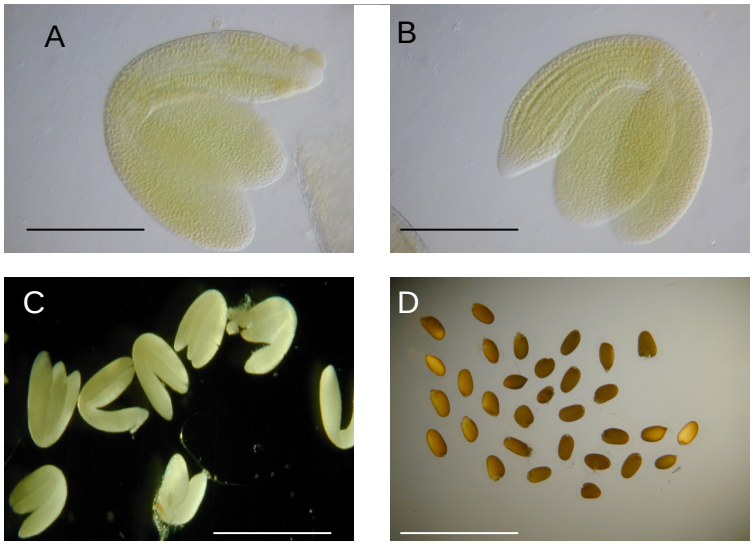

Fig. 3. 1. Morphology of *swn-7 clf-28* mutant embryos.

A selection of cleared embryos from self pollinating *swn-7/-7 clf-28/+* segregating plants; ¼ being *swn-7/-7 clf-28/-28*. (A) An example of a wild-type excised embryos after 1 day imbibition. (B) An example of excised progeny of *swn-7/-7 clf-28/+* embryo (all progeny showed the wild-type phenotype after 1 day imbibition). (C) A selection of excised embryos from *swn-7/-7 clf-28/+* siliques. (D) seeds from *swn-7/-7 clf-28/+* plant. No morphological differences were observed. (Scale: A and B = 0.3 mm, C = 7 mm, and D = 15 mm)

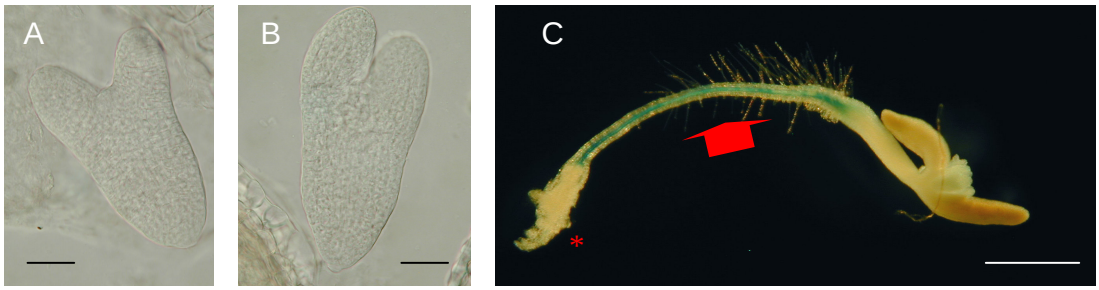

Fig. 3. 2. Expression of *AP1::GUS* in *swn- clf-* embryos.
(A) Wild-type *AP1::GUS* embryo. (B) An example of an embryo from siliques of *swn-7/-7 clf-28/+ AP1::GUS* plants (segregating ¼ *swn-7/-7 clf-28/-28*) were stained at various developmental ages, and no expression was observed in any of the embryos, suggesting that *AP1::GUS* is not mis-expressed in *swn-7 clf-28* embryos. (C) *swn-7/-7 clf-28/-28 AP1::GUS* double mutants at two weeks old showing ectopic expression in "root" material (arrowed) but not in the "pickle" root tip (asterisk) (image kindly donated by Dr. O. Clarenz). (Scale: A and B = 0.2 mm, C = 2 mm)

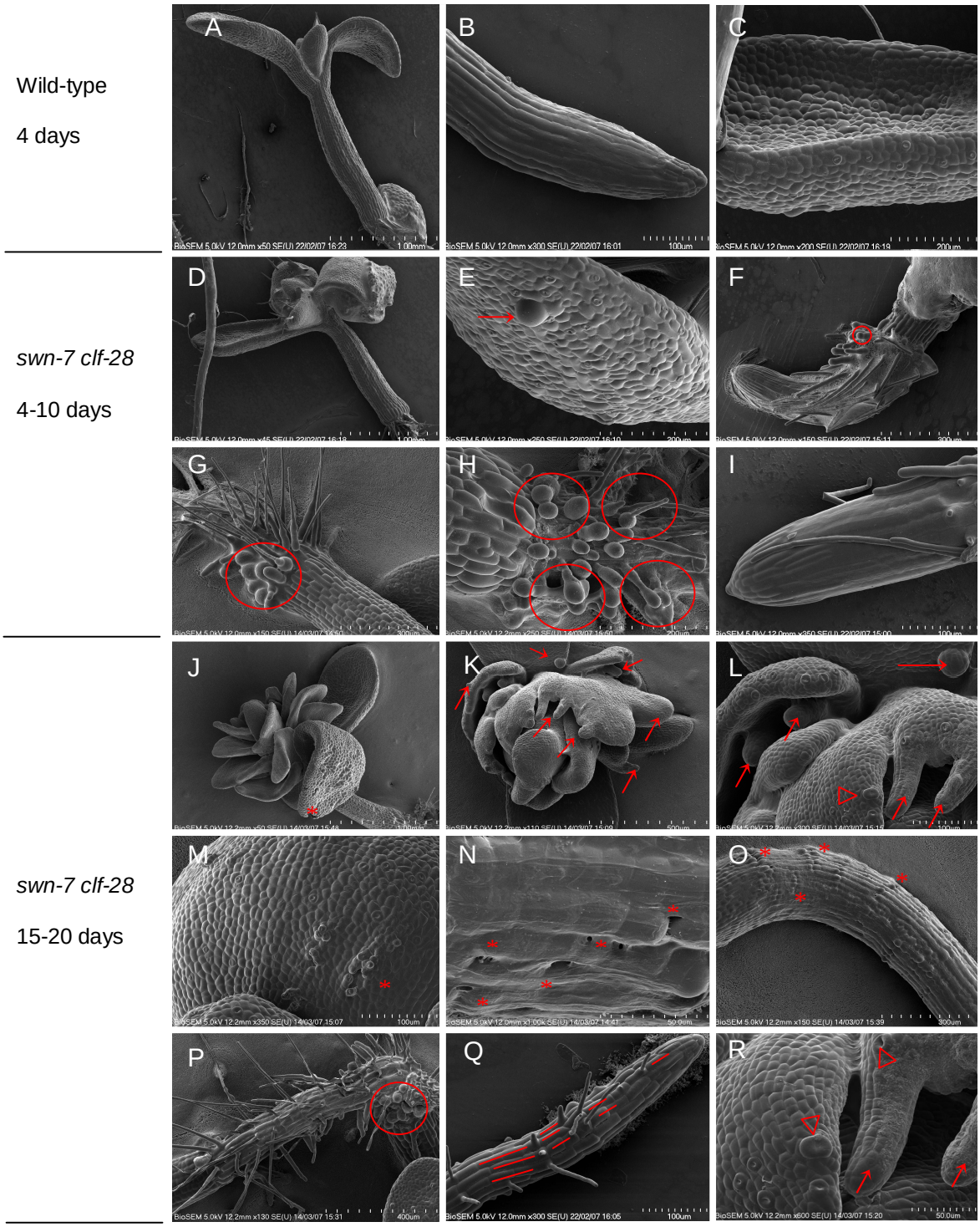
Wild-type
4 days
A
B
C
swn-7 clf-28
4-10 days
D
E
F
G
H
I
swn-7 clf-28
15-20 days
J
K
L
M
N
O
P
Q
R

Fig. 3. 3. Scanning electron micrographs comparing morphology wild-type and *swn-7 clf-28* seedlings.

4 day old **wild-type** seedlings (A-C): A = aerial parts, B = root tip, and C = cotyledon surface. 4-10 days ***swn-7 clf-28* mutants** (D-I): D = aerial parts, E = cotyledon surface (arrow indicates irregular leaf outgrowths), F = whole root, G and H = root-hypocotyl junction (circles in F, G + H indicate root hair defects), and root tip (I). 15-20 day old *swn-7 clf-28* (J-R): J, K, L, M + R = aerial parts (asterisk indicates cell adhesion defects, arrows indicate aberrant outgrowths, and open triangle indicates stomata defects). N, O, P, +Q = primary root (asterisk indicates cell adhesion defects, arrows indicate aberrant outgrowths, lines highlights epidermal cell files demonstrating root hair occurring in contiguous files, and open triangle indicates stomata defects). The deformities identified in the *swn-7 clf-28* mutants were identical to *vrn2-1 emf2-3* mutants (n = ~10 plants per time point)

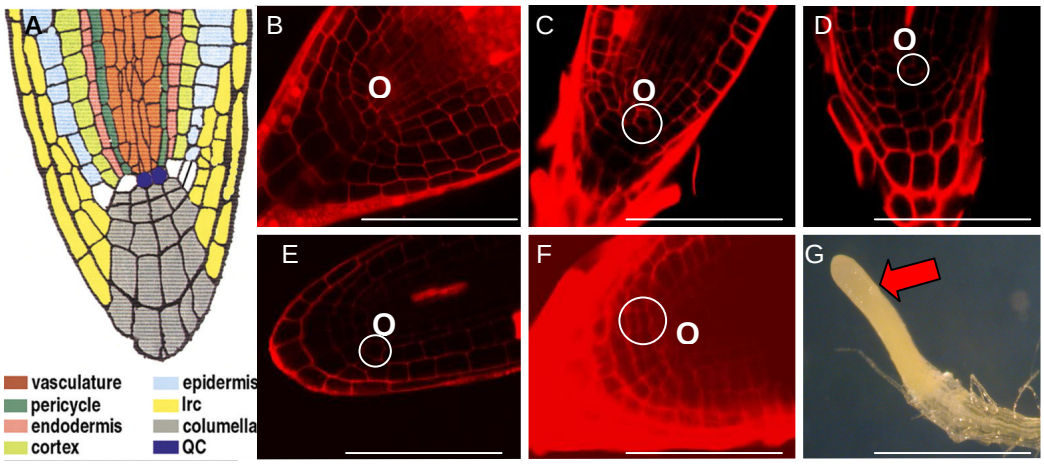

Fig. 3. 4. Internal root architecture of wild-type, *swn- clf-* and *vrn2-1 emf2-3* root tips.

After 4 days, irregular cell size, orientation, and shape in the *swn-1 clf-50*, *swn-7 clf-81*, *swn-7 clf-28*, and *vrn2-1 emf2-3* was compared to wild-type roots. (A) Schematic of the *Arabidopsis* root tip (Mylona *et al.*, 2002) (B) wild-type, (C) *swn-1 clf-50*, (D) *swn-7 clf-81*, (E) *vrn2-1 emf2-3*, (F) *swn-7 clf-28* (with transformed root tip), and (G) *swn- clf-28* ("pickle" root tip). Arrow indicates transformed root tip, circles indicated cell morphology defects, and "Q" denotes the QC. (Scales: B -F = 1 mm, G = 5 mm)

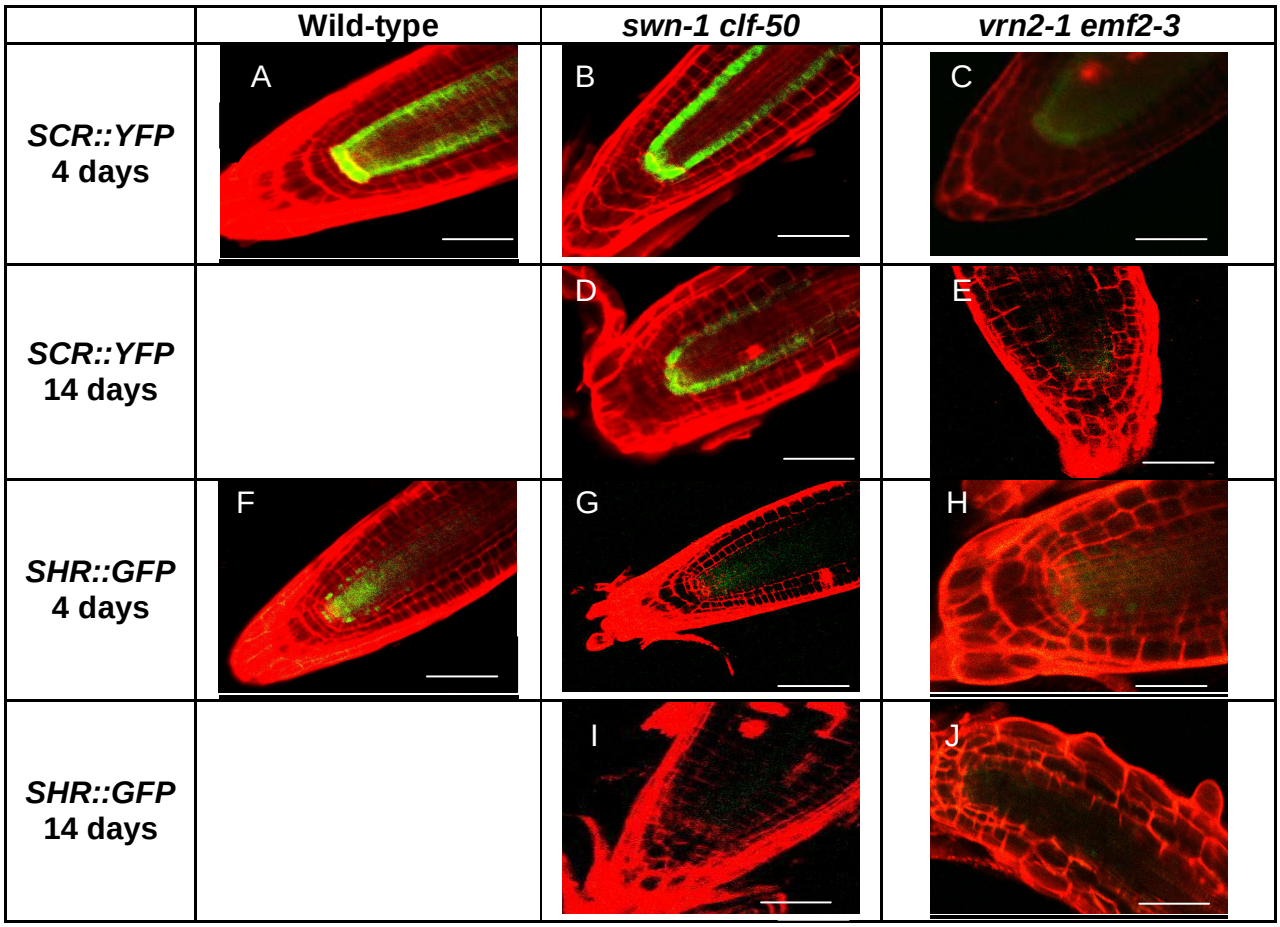

Fig. 3. 5. The effect of *swn- clf-*, and *vrn2-1 emf2-3* mutation on *SCR::YFP* and *SHR::GFP* expression.

Samples were analyzed after 4 and 14 days growth, no observable difference between wild-type samples and double mutants was consistently observed after 4 days. After 14 days *swn-1 clf-50*, *swn-7 clf-81*, and *vrn2-1 emf2-3* lines showed reduced expression of *SCR::YFP*, *SHR::GFP (swn-7 clf-81* data (not shown) is identical to *vrn2-1 emf2-3* (Dr. B. Scheres). (Scale = 1 mm). Wild-type: (A) wild-type *SCR::YFP* 4 days, (F) *SHR::GFP* 4 days (C). *swn-1 clf-50: SCR::YFP* 4 and 14 days, (B + D) *SHR::GFP* 4 and 14 days (G + I). *vrn2-1 emf2-3: SCR::YFP* 4 and 14 day days (C+ E), *SHR::GFP* 4 and 14 days (H + J)

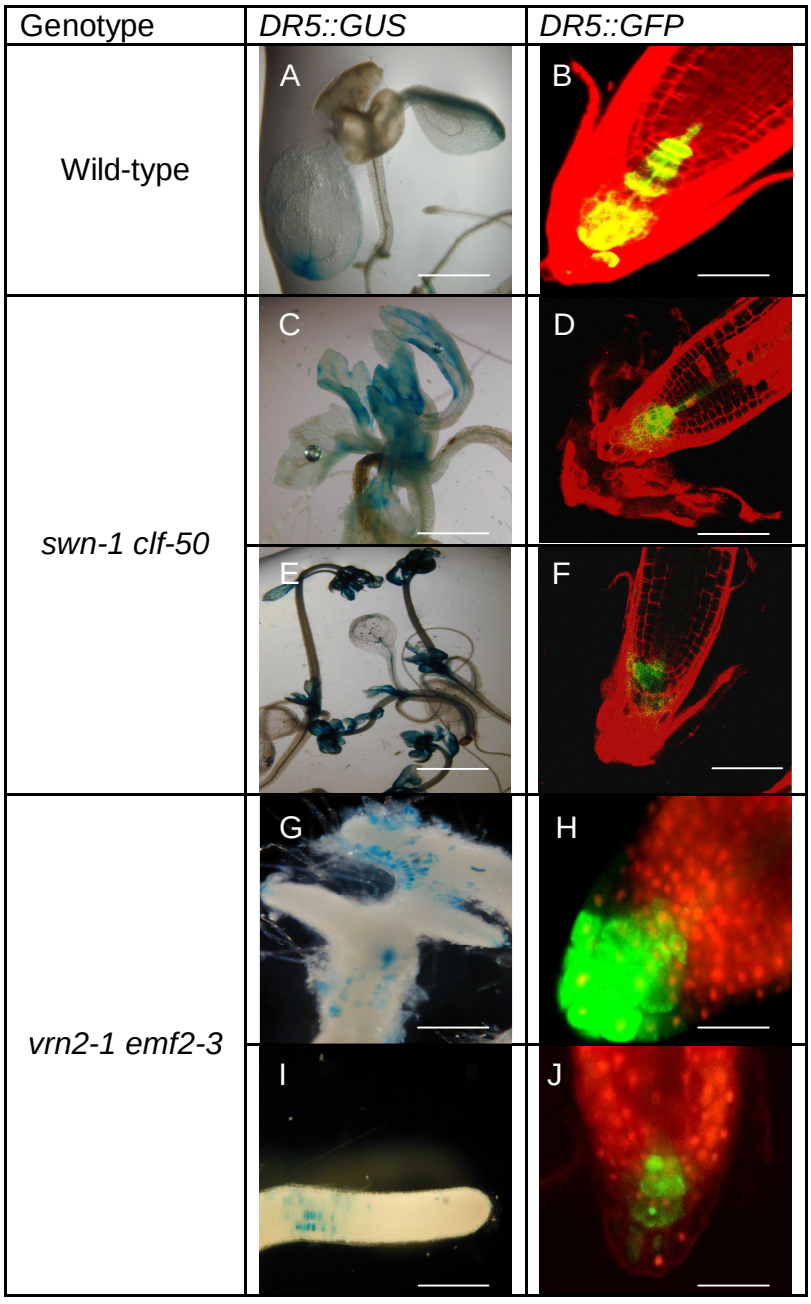

Fig. 3. 6. Auxin distribution in *swn-clf-* and *vrn2-1 emf2-3* double mutants.
Expression *DR5::GUS* in 10 days wild-type is seen in root tips and tips of leaves and cotyledons (A) expression is detect in leaf tips, (B) *DR5::GFP* expression in wild-type root tips.

DR5::GUS expression in 14 day old *swn-1 clf-50* seedlings (C + E), expression is observed throughout leaves and floral organs. *DR5::GFP* at 4 days in *swn-1 clf-50* mutants (D) and 14 days (F) expression is reduced in *swn-1 clf-50* root tips after 14 days. *DR5::GUS* expression in *vrn2-1 emf2-3* at 14 days (G + I). *vrn2-1 emf2-3*, showing abnormal auxin distribution. *DR5::GUS* 4 (H) and 14 days (J) in *vrn2-1 emf2-3* mutants showing reduced expression in the root tip after 14 days. Pc-G mutants show decreased auxin accumulation in the root tip and increased accumulation in the aerial parts and callus-like tissues. (Scale: A, C, E G, I = 10 mm, and B, D, F, H, J = 1 mm)

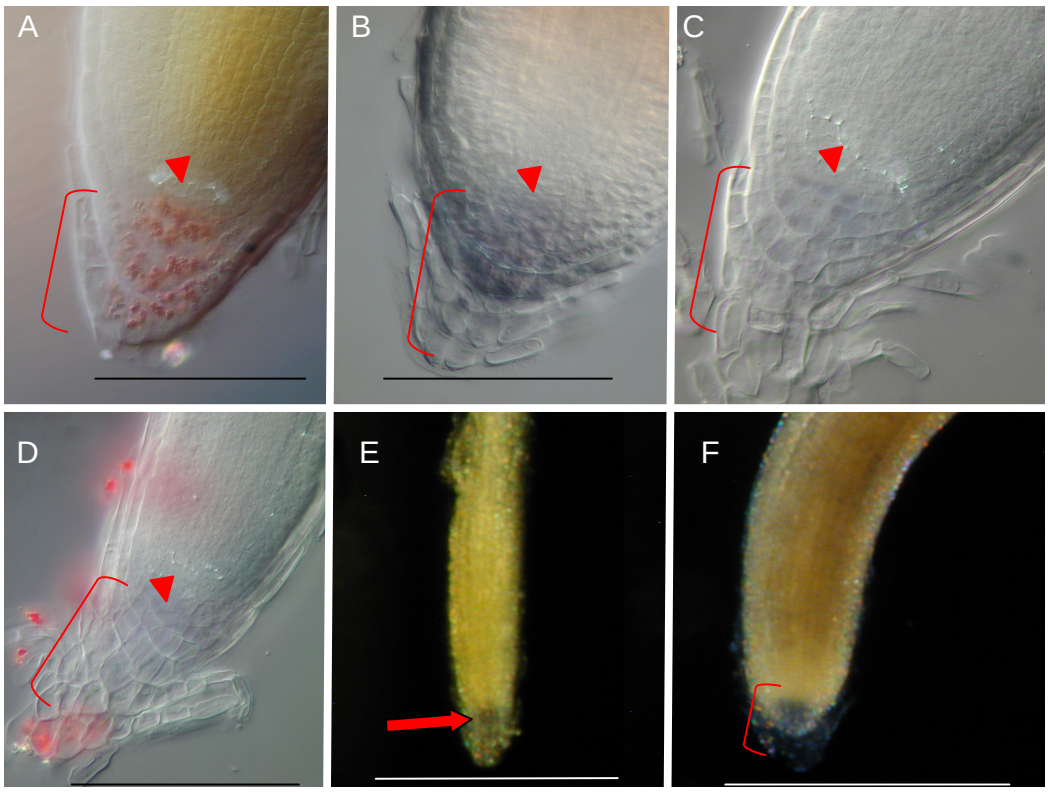

Fig. 3. 7. The effect of *swn- clf-* and *vrn2- emf2-3* mutants on differentiation in the root tip.

4 day old *swn-7 clf-81*, *vrn2-1 emf2-3* and wild-type plants seedlings stained for starch using Lugol's stain of. Wild-type (A and E), *swn-7 clf-81* (B), *swn-7 clf-28* (C + F), and *vrn2-1 emf2-3* (D). The *swn- clf-* and *vrn2-1 emf2-3* seedlings show an expanded zone of staining in the root tip. Arrows and brackets indicate starch staining, and triangle indicates the location of the QC (Scale = A-D = 1 mm, E + F = 5 mm)

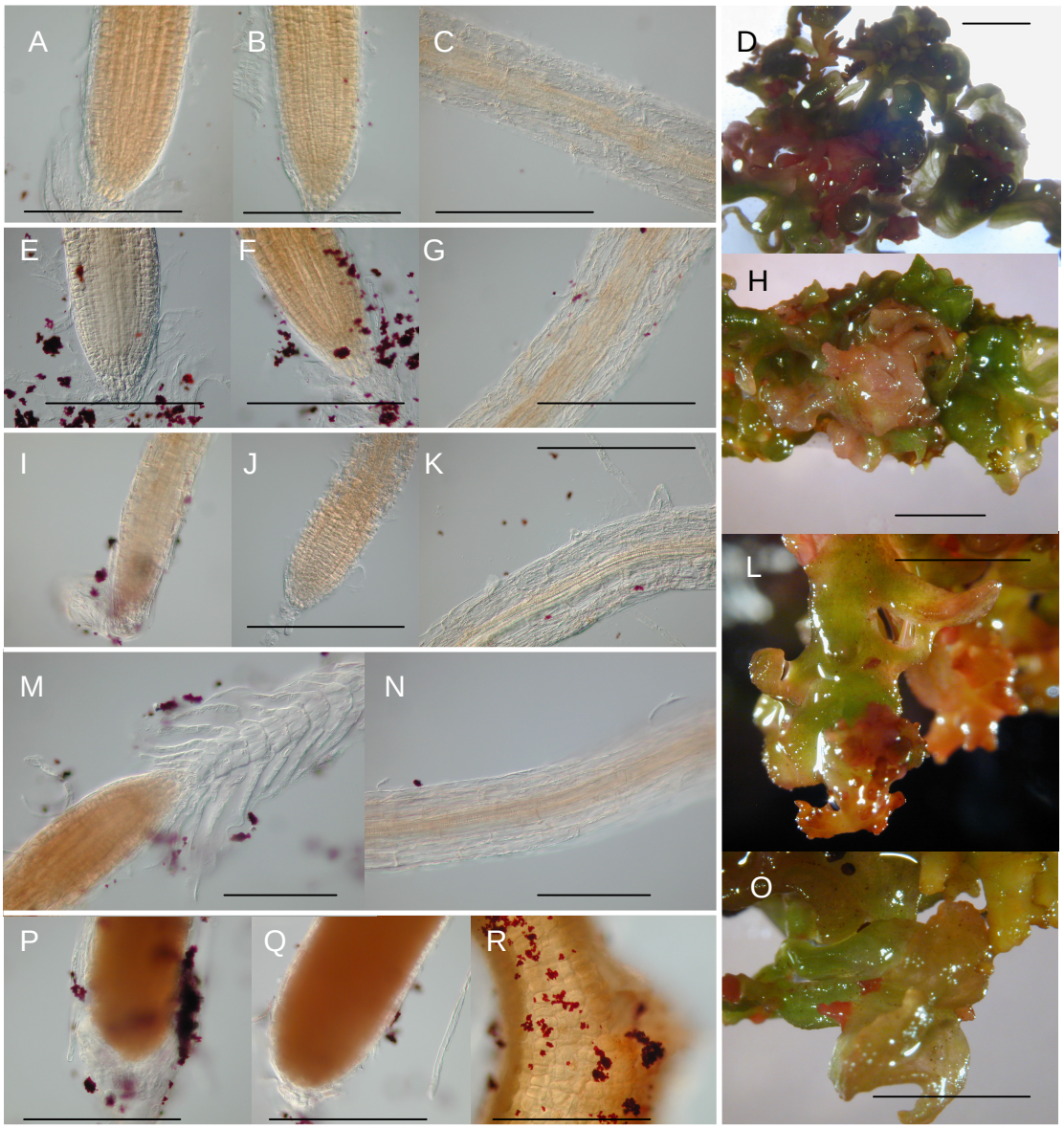

Fig. 3. 8. Staining for embryonic storage oils in *swn- clf-* and *vrn2-1 emf2-3* mutants.

Mutant and wild-type plants were stained with Sudan 7 and evaluated for accumulation of embryonic storage oils in the root tissues. No difference was found between *swn-1 clf-50* (E-G), *vrn2-1 emf2-3* (I-K), *swn-7 clf-81* (M + N), *swn-7 clf-28*, *swn-7 clf-81*(P-R), and wild-type (A-C) plants at 4 or 14 days. The *swn-7 clf-81* callus-like tissue showed Sudan 7 staining indicating the presence of embryonic storage oils (D, H, L, O). Dark red granules are residual, undissolved Sudan 7 stain that remained after washing. (Scale = 2 mm, n = 5-10 for each genotype)

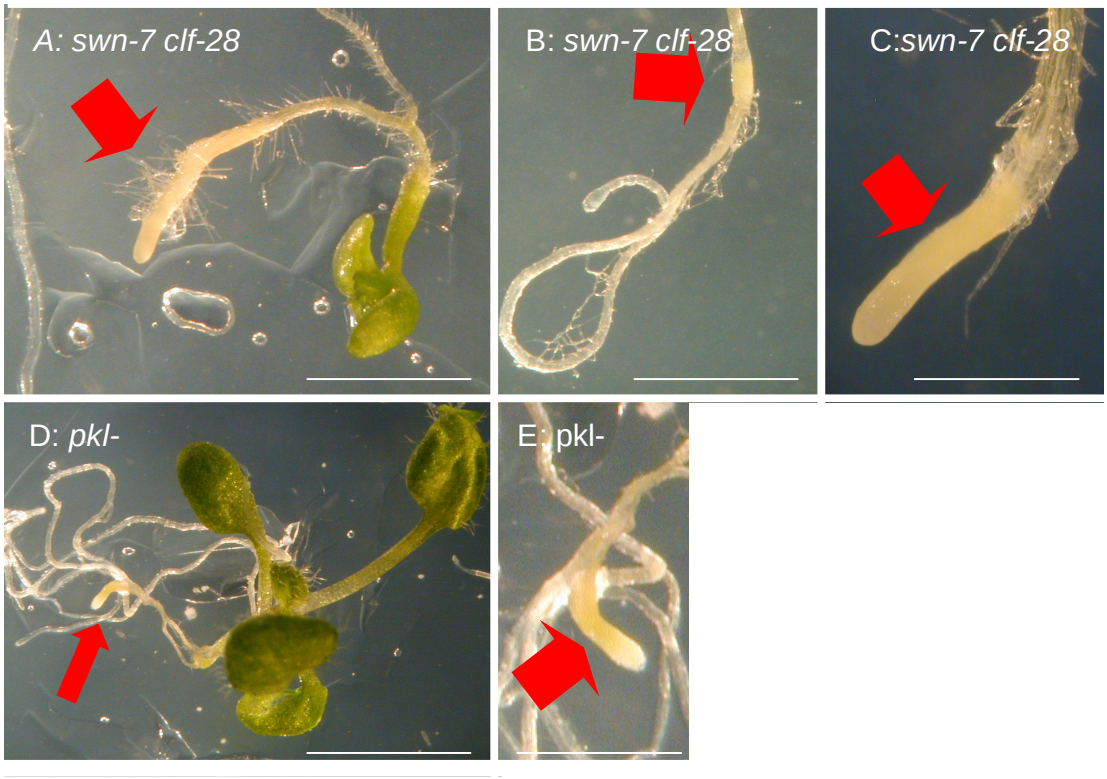

Fig. 3. 9. Pickled root phenotype of *pkl*- and *swn- clf*-mutants.
10 day old *swn-7 clf-28* and *pkl-1* seedlings. (A-C) *swn-7 clf-28* seedlings with transformed root tips. (B) Transformation of *swn-7 clf-28* plants does not always occur at the tip. (D + E) *pkl-1* seedling root showing transformation always at the root tip. Red arrows indicate the transformation. (Scale: A, B, D =10 mm, B + E = 2 mm).

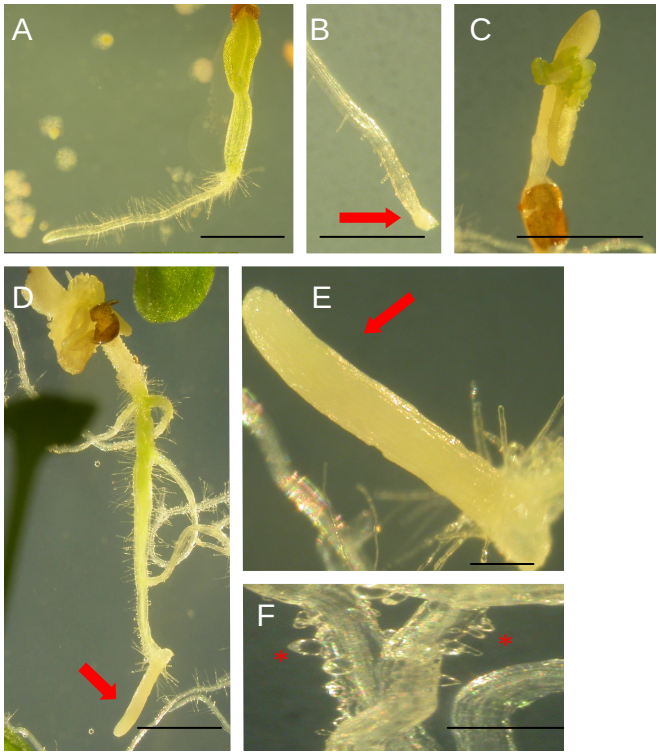

Fig. 3. 10. The effect of sugar concentration on *swn-2 clf-81* phenotype.
swn-2 clf-81 mutants grown on 0% (A), 1% (B + C), and 4% (E- I) sugar after 10 days. When grown on 4% sucrose there is an increase in transformed roots and root hair defects, the aerial parts were seemingly unaffected. The *swn-7 clf-81* phenotype was variable. Arrows indicate transformation, asterix denotes root hair deformity. (Scale: A = 5 mm, B, C, E and F = 1 mm, D = 2 mm)

4. 0. Analysis of global gene expression of *swn-7, clf-28,* and *swn-7 clf-28* seedlings

Only eight Pc-G targets have been so far validated, it is likely many more targets exist considering the gross phenotypic deformities of the Pc-G mutants. The known Pc-G targets are transcription factors, therefore the targets of the Pc-G are likely to have downstream effects on transcription. *SWN* and *CLF* are likely to act redundantly to catalyze the H3K27me^3 mark which correlates with gene repression or silent chromatin in order to regulate proper development (Schubert *et al.*, 2006; Lindroth *et al.*, 2004). Therefore, genes that possess H3K27me^3 in wild-type plants are likely direct Pc-G targets, and the loss of this repressive mark in *swn-7*, *clf-28*, and *swn-7 clf-28* mutants is likely to cause their de-repression.

Transcriptional profiling of wild-type, *swn-7*, *clf-28*, and *swn-7 clf-28* seedling was conducted with the following aims: Firstly, to gain insight into the mutant phenotypes of *swn-, clf-, swn- clf-* plants by identifying the genes mis-expressed in the different mutants backgrounds, particularly of genes mis-expressed in *swn-* that are associated with the J-A transition. Secondly, to identify potentially novel direct targets by comparing genes mis-expressed in Pc-G mutants to the genome wide H3K27me^3 map (Zhang *et al.*, 2007). The genes mis-regulated in *swn-7*, *clf-28*, and *swn-7 clf-28* mutants that possess the H3K27me^3 mark are likely direct Pc-G targets, whereas, genes that lack H3K27me^3 and are mis-expressed are likely to be secondary targets. Thirdly, to resolve the proposed redundancy of *CLF* and *SWN*, and their discrete functions by comparing the transcription profiles of *swn-*, *clf-*, and *swn- clf-* mutants.

Affymetrix Genechip array (ATH1) was under-taken by the NASC's International Affymetrix Service on two biological replicates of RNA, extracted using a Qiagen RNeasy kit, from *swn-7*, *clf-28,* and wild-type (Col) 10 days old seedlings, and *swn-7 clf-28* at 12 days old. The ATH1 contained 22,500 probe sets of 25-olimers in length, representing approximately 24,000 genes. There are less probe-sets than represented genes due to probe promiscuity. Previous investigations had shown that the

developmental defects of *swn-7 clf-28* seedlings increased in severity over time (Chanvivattana *et al.*, 2004). This is likely to correlate with increased mis-expression of secondary targets, caused by the mis-regulation of direct targets. To increase the likelihood of finding direct target genes of *SWN* and *CLF* 10 day old seedlings were analyzed. The *swn-7 clf-28* mutants exhibit slow growth, and so to allow a more appropriate comparison of this developmental stage *swn-7 clf-28* were grown for an extra two days compared to the other genotypes. In addition, the *swn-7 clf-28* double mutants are sterile and so *swn-7/-7 clf-28/28* mutants were selected from the progeny of a self-pollinated *swn-7/-7 clf-28/+* plants. The progeny were grown on 0.5% MS plates and segregate in a 1:4 ratio of double mutants, and are difficult to unambiguously select from wild-type seedlings until ~10 days. The *swn-7, clf-28,* and *swn-7 clf-28* alleles were analyzed as both alleles are null mutation (Personal communication Dr. W. J. Goodrich), and the *swn-7 clf-28* mutant seedlings showed the strongest phenotype compared to other lines in respect to morphology and root transformation penetrance.

Background correction, normalization, and gene expression analysis of the array data were performed using the GC-RMA routine in GeneSpring version 7.2 (Silicon Genetics). Absent signals where removed from the data set and averaged signals of two biological replicates were calculated and the *swn-7, clf-28, swn-7 clf-28* expression profiles were compared to the wild-type. Signals that were at least two-fold different to wild-type were considered mis-expressed in the mutant genotypes. The "Sungear" and "biomap" functions on www.virtualplant.org websites were used to compare gene sets and their biological relevance throughout this chapter.

4. 1. Validation of microarray data

As a preliminary step to validating the microarray data, I checked whether previously identified targets were well represented. Prior to this study only eight targets post-germination had been validated. *CLF* targets included *AG* and *AP1,* which are found ectopically expressed in *clf-* plants (Chanvivattana *et al.,* 2004; Goodrich *et al.,* 1997) . CLF is found bound to the *AG* loci (Schubert *et al.,* 2006), and *AG* and *AP1* are known

to possess H3K27me^3 mark (Schubert *et al.*, 2006; Zhang *et al.*, 2007). *SWN* and *CLF* act redundantly to repress *PHE, MEA, FUS3, STM, AGL19,* and *FLC* (in a vernalization-requiring background). These genes were identified as targets of the Pc-G by their ectopic expression in *swn- clf-* mutants and they are known to possess the H3K27me^3 mark (Makarevich *et al.*, 2006; Katz *et al.*, 2004; Wood *et al.*, 2006; Chanvivattana *et al.*, 2004; Schonrock *et al.*, 2006; Jullien *et al.*, 2006). CLF is found bound to the *FLC, FUS3, AGL19,* and the *STM* loci (Schonrock *et al.*, 2006; Makarevich *et al.*, 2006; Wood *et al.*, 2006), but both SWN or CLF are required for their repression. Most of these genes were found to be up regulated in the array data. Strikingly, in the *clf-28* mutants *AG* is the most highly up regulated gene being 50 fold up regulated, and *AP1* is also highly up regulated (appendix: Table. 8. 1). For the *swn- clf-* double mutant known redundant *SWN CLF* targets were found highly up regulated, such as *STM*, *FLC*, and *FUS3* (appendix: Table. 8. 1). However, several other known targets were not detected in the *swn-7 clf-28* data set, for example, *AGAMOUS-LIKE 19* (*AGL19*), *MEA*, and *PHE1,* even through probe sets specific for these genes are found in the Affymetrix Genechip. This may be due to the microarray not being sensitive enough to detect their changed expression, and this may be a consequence of as ectopic expression occurring in cell types restricted to small regions such as the quiescent centre in the root, or a specific developmental stage for example in flowers which are not formed in *swn- clf-* mutants. Whether this is the case for *AGL19* is unclear, as previous research that discovered *AGL19* as a target of Pc-G repression was carried out on fully whole plants that were flowering, and in this piece of work, the analysis of *AGL19* expression pattern was not carried out on the *swn- clf-* mutants. Alternatively, the mis-expression of these genes may occur in tissues not found in the early seedling stages of the *swn- clf-* mutants but may be expressed in the callus-like material found later. Identifying known Pc-G targets in the array data partially validates the effectiveness of the microarray data and indicates that the microarray analysis can identify most, but not all, genes regulated by the Pc-G. This suggests that the datasets may be used to identify potential novel direct targets of the Pc-G (as described later).

In order to further test the legitimacy of the microarray data, RT-Q-PCR and RT-PCR was carried-out using independent biological replicates distinct from the samples used for microarray analysis. The analysis was carried out on mis-regulated genes expected to be direct targets, firstly for the *swn-7* data and then the *swn-7 clf-28* and *clf-28*. For the *swn*-7 data, I prepared cDNA from two independent *swn*- alleles (*swn-2* and *swn-7*) at 10 days old. I first tested At5g15160 (*similar to PACLOBRUTRAZOL RESISTANT1*) which was one of the most highly up regulated genes in the *swn-7* array (appendix: Table. 8. 1) and is found to possess H3K27me^3 (Zhang *et al*., 2007) a strong indication that it is a direct SWN target. Figure 4. 1 shows that At5g15160 is up regulated in *swn-2* and *swn-7*. This indicates that it is repressed by the SWN-Pc-G, as it is found up-regulated in two independent alleles. I then tested the expression of two auxin-responsive family proteins (At3g53250 and IAA29), a short-chain dehydrogenase/reductase (SDR) protein (At5g02540), a GA-20-oxidase (At5g51810), a zinc ion binding protein (At5g6110), and an oxidoreductase protein (At5g22500). Their expression was found not to exceed two fold up regulation in both *swn-2* and *swn-7* alleles compared to wild-type seedlings (data not shown), indicating that they may not be direct targets of *SWN* or the nature of the mutations can cause differential expression. The microarray data suggested *similar to PACLOBRUTRAZOL RESISTANT1* was up regulated in *swn-7* by >7 fold, but the QPCR analysis showed only 2-3 fold increase. This may indicate that fewer genes are up regulated in *swn-7* than the array data suggests.

For *swn-7 clf-28*, I selected genes found highly up regulated and possessed H3K27me^3, so represent potential direct targets of the Pc-G, these included: *LEC1*, *FUS3, ABI3, ABI4, EM1*, *CUP SHAPED COTYLEONS* (*CUC2*), *LATERAL ORGAN BOUNDARIES 16* (*LBD16*), *LBD40*, *ANTEGUMENT-LIKE* 5 (*AIL5*), *AIL7, GIBBERRELIC ACID 2-OXIDASE* (*GA2-OX*), *YUCCA4* (*YUC4*), *AGL11, AGL14, AGL16, AGL67* and *TRANSPARENT TESTA16* (*TT16*). The ectopic expression of these genes may also play a causative role in the *swn- clf-* phenotype as they predominantly represent genes of developmental importance, such as SAM maintenance, lateral organ formation, hormone synthesis, and regulators of embryonic traits, processes of which have been implicated in

being disrupted in *swn- clf-* mutants (Chapter 3). Figure 4. 2 shows that these genes are indeed up regulated in *swn-7 clf-28* mutants. I also tested *FLC* and *FT* as they are key regulators of flowering time and the Pc-G mutants are shown to exhibit flowering time defects, *FLC* was found up regulated and *FT* down regulated in *swn-7 clf-28* (Fig. 4. 2) as is consistent with the microarray results. *FLC* was found up regulated in the *clf-28* microarray data set and this was validated as being up regulated in independent *clf-28* and *clf-81* alleles by Q-PCR (Fig. 4. 3). Together these results suggest: (A) that known Pc-G targets are well represented, but not all are found mis-expressed, (B) The *swn-7* array data may not be very robust as only 1 of 7 genes could be validated, and (C) The *swn-7 clf-28* array data appears robust and reproducible, at least for genes that are highly mis-expressed.

4. 2. The microarray data suggests *CLF* and *SWN* act redundantly at most targets

770 genes are mis-expressed in *swn-7* seedlings with a two-fold expression difference relative to wild-type expression (Fig. 4. 4 and appendix: Table. 8. 1 and Table. 8. 2), a large number considering that *swn-* plants have no gross morphological defects. Nevertheless, as shown in Chapter 2, *swn-* shows subtle phenotypic differences to wild-type plants, and gene mis-expression found here may be responsible for this. Another explanation is that the mis-expressed genes have no effect on morphology, or, the differences found in expression are inaccurate.

There are 797 genes mis-expressed in *clf-28* seedlings with at least a two-fold expression difference to wild-type expression (Fig. 4. 4 and appendix: Table. 8. 1 and Table. 8. 2). A large degree of mis-expression was expected considering that *clf-* seedlings have visible phenotypic deformities (Kim *et al.*, 1998; Puangsomlee, 1997; Goodrich *et al.*, 1997; Chanvivattana *et al.*, 2004). Surprisingly, the number of genes mis-expressed in *clf-28* seedlings is only marginally more than found in *swn-7*. However, the scale of mis-regulation in *clf-28* is much greater than *swn-7*, as mis-expressed genes in *clf-28* range from 50 fold up to 33 fold down. Of the genes up regulated 288/361 are between 2-3.3 fold. Whereas, *swn-7* mutants displayed mis-expression ranging from 16 fold up and 25

fold down regulated, and 240/255 genes up regulated are only by 2.5 fold (appendix: Table. 8. 2). This may explain why many of the target genes were not able to be validated for *swn-7,* as it is likely much of the mis-expression is "noise" or background signal rather than being a consequence of being mis-expressed due to the *swn*- mutation.

6030 genes (22.3% of the *Arabidopsis* transcriptome) are mis-expressed in *swn-7 clf-28* seedlings (Fig. 4. 4 and appendix: Table. 8. 1 and Table. 8. 2). The scale of gene mis-expression demonstrates the crucial importance of the Pc-G, and the complicated nature of investigating such a keystone in developmental regulation. The extensive mis-regulation in *swn-7 clf-28* compared to the single mutants supports the hypothesis that *SWN* and *CLF* act redundantly as is also suggested by the severity of the double mutant phenotype relative to the single mutants.

swn-7 has 131 uniquely up regulated genes (i.e. only in *swn-7*), and *clf-28* has 239 genes uniquely up regulated (Fig. 4. 4). *swn-7* and *clf-28* have 119 genes commonly up regulated (Fig. 4. 4), meaning they are up regulated in both mutant genotypes. *swn-7 clf-28* has 2566 genes that are only up regulated in this genotype; this is likely the result of *SWN* and *CLF* functional redundancy (Fig. 4. 4). This pattern is replicated in the genes down regulated, with *swn-7* showing 129 genes and *clf-28* 148 discretely down regulated, 291 genes are found down regulated in both *swn-7* and *clf-28*. 2976 genes are found to be down regulated due to the redundant function of *SWN* and *CLF* (Fig. 4. 5).

Thus, the analysis of genes up regulated in the different mutants revealed that *SWN* and *CLF* carry out discrete, common, and redundant roles, i.e. the regulation of genes can require either *SWN* or *CLF* (so called discrete targets), which are mis-regulated in either of the single mutants. Other loci require both *SWN* and *CLF* for regulation (so called common targets) and are mis-regulated in both of the single mutants. Other loci require either *SWN* or *CLF* to regulate expression correctly (redundant targets) and are only mis-regulated in the double mutant.

The severe and pleiotropic phenotype and the number of genes mis-expressed in *swn-7 clf-28* mutants compared to *swn-7* or *clf-28* further supports the hypothesis of *SWN* and *CLF* functioning redundantly. Whereas the number of genes mis-expressed in *swn-7* and *clf-28* mutants is similar, which is surprising considering the phenotypic differences, closer inspection reveals that it is type of genes specifically mis-expressed in *clf-28,* and not in *swn-7*, predominantly MADs-box transcription factors, are the likely causes of the *clf-* phenotypes. As the genes mis-expressed are both direct and indirect target of the Pc-G, further analysis was carried to uncover which genes found mis-expressed are direct P-G targets.

4. 3. Identifying potential direct Pc-G targets, by comparing mis-expressed genes in Pc-G mutants to the H3K27me^3 epigenome map

Genes that were found to be mis-regulated in *swn-7, clf-28* and *swn-7 clf-28* mutants, and overlaps between them, were compared to genes known to possess H3K27me^3 in wild-type seedlings. This was carried out to find potential direct targets of *SWN* and/or *CLF*. As *SWN* and *CLF* are known to be required to create/maintain this mark to regulate gene expression (Lindroth *et al.*, 2004; Schubert *et al.*, 2006; Chanvivattana *et al.*, 2004; Wood *et al.*, 2006). No other proteins have been found that can make the H3K27me^3 mark in wild-type plants. The mis-expression of genes found in *swn-7, clf-28* and *swn-7 clf-28* mutants that possess the H3K27me^3 mark in wild-type plants would probably represent direct targets of the Pc-G, and those genes lacking the H3K27me^3 mark that are mis-expressed are the likely consequence of secondary effects, so called indirect targets.

A previous whole genome "ChIP on chip" study suggested that 4596 loci possess H3K27me^3 corresponding to ~17% of the genome (Zhang *et al.*, 2007). These results are consistent with an independent study of genes possessing H3K27me^3 on chromosome 4 (Turck *et al.*, 2007). 21.8% (55) genes up regulated in *swn-7* also possess H3K27me^3 (Fig. 4. 6 and Table. 4. 1) Down regulated genes in *swn-7* shows 25.5% (135) genes possess H3K27me^3 (Fig. 4. 7 and Table. 4. 1) . This illustrates there are 591 (75.6%)

mis-expressed genes that do not possess H3K27me^3, and these are considered as secondary targets. This indicates that the majority of mis-expression is due to indirect effects. *swn-7* has a marginally higher percentage of genes that possess H3K27me^3 that are up or down regulated than the genome frequency of H3K27me^3, which may indicate that the there is weak correlation of ectopic expression in *swn-7* and the H3K27me^3 mark. These genes represent potential discrete direct targets of *SWN*.

Analysis of genes up regulated in *clf-28* shows that 31.1% (112) genes possess H3K27me^3 (Fig. 4. 6 and Table. 4. 1), this is more than double the number of direct targets found up regulated in *swn-7*. Down regulated genes show 23.6 % (104) genes possess H3K27me^3 (Fig. 4. 7 and Table. 4. 1). Therefore, 588 (73.1%) mis-expressed genes are indirect targets. 20.6% (557) genes possess H3K27me^3 that are up regulated in *swn-7 clf-28* (Fig. 4. 6 and Table. 4. 1). 17.1% (576) genes possess H3K27me^3 that are down regulated in *swn-7 clf-28* mutants (Fig. 4. 7 and Table. 4. 1). Thus, there are 4944 (81.9%) genes mis-expressed as indirect targets. There is a higher percentage, and more genes, up regulated that possess H3K27me^3 than down regulated, consistent with the role of the Pc-G as a repressor complex.

SWN and *CLF* have a potential of 1004 direct redundant targets i.e. genes that are only up regulated in *swn-7 clf-28* and possess H3K27me^3. It can be also be concluded that *CLF* has 129 potential discrete direct targets. *SWN* and *CLF* have 91 potential target genes in common, where both *SWN* and *CLF* are required for proper regulation. (Fig. 4. 6 and Fig. 4. 7).

There are many potential direct target of the Pc-G (1333 genes in total (total number of gene up regulated gene in *swn-7, clf-28, swn-7 clf-28* and commonly up regulated genes)), many more than is found in flies (100-200 direct targets (Ringrose *et al.*, 2003; Schwartz *et al.*, 2006)). However, there are ~4600 genes that possess H3K27me^3 in the *Arabidopsis* genome, so the majority of H3K27me^3 carrying genes were not detectably mis-expressed in *swn-7 clf-28* mutants, which lack Pc-G activity. There are several obvious explanations why mis-expression was not found: A) subtle mis-expression in

the mutants is below the two fold threshold of the array analysis, for example if a gene is ectopically expressed in relatively few cells. B) The Pc-G may repress these genes tissue specifically, for example repressing gene expression in leaf tissue, but the *swn-7 clf-28* mutants lack fully developed leaves so their mis-expression may not be detected. C) H3K27me^3 possessing genes may require specific developmental or environmental signals for their expression, and without them are not expressed. These signals may not have occurred in 10/12 day old seedlings, thus their mis-expression may not have occurred in the mutants in the conditions tested. Alternatively, not all of the H3K27me^3 possessing genes are Pc-G targets. This might occur if H3K27me^3 is catalyzed by other SET domain proteins. Yet another alternative is that these genes may be H3K27me^3 marked by the Pc-G but removing Pc-G function may not cause their mis-expression because other factors, perhaps histone marks, may be required to co-ordinate gene expression which are not present in *swn-7 clf-28* mutants.

The percentage of genes mis-regulated that are direct targets is relatively low in each of the genotypes analyzed indicating that most of the transcriptional changes are in fact an indirect consequence of mis-expression of the direct targets. This appears particularly true of the *swn-7 clf-28* mutants, which may indicate that *SWN CLF* act redundantly to regulate particularly promiscuous transcription factors, perhaps master regulators of transcriptional profiles.

The percentage of genes in *clf-28* that possess H3K27me^3 is nearly twice the genome average, supporting the finding that *CLF* mediated H3K27me^3 mark and gene repression are functionally linked. Previous investigations of the Pc-G has demonstrated its role in transcriptional repression (Chanvivattana *et al.*, 2004; Goodrich *et al.*, 1997; Schubert *et al.*, 2006). The data presented here shows virtually half of the genes mis-regulated in each of the mutants analyzed are down regulated.

4. 4. The H3K27me^3 decorated genes are most strongly mis-expressed in Pc-G mutants

It was noted that many of the known direct Pc-G targets were amongst the highest up regulated genes. For example, *AG* was the most strongly up regulated gene in *clf-28,* and *FUS3* was amongst the most highly mis-expressed in *swn-7 clf-28* (appendix: Table. 8. 1). To test whether this was generally true, mis-expressed genes were ranked according to their level of up regulation, these were then partitioned in to groups with a similar level of mis-expression, i.e. the top 5% most highly up regulated, then the top 10%, 10-20%, 20-50%, and the bottom 50% of genes up regulated. Within each group the proportion of genes possessing H3K27me^3 were analyzed.

Table. 4. 2 shows that the most highly up regulated genes in *swn-7, clf-28,* and *swn-7 clf-28* have a much higher proportion of genes that possess H3K27me^3 than the genome average. For example, *swn-7 clf-28* has overall 20.6% (557 genes) of its total up regulated genes possessing H3K27me^3 only marginally above the genome average of 17%, but this is massively enriched when analysing the top 10% (274) of genes up regulated, which has over 57% (156) of genes possessing H3K27me^3. This pattern appears to act on a sliding scale with the highest percentage of H3K27me^3 marked genes being found in the most highly up regulated, this falls when looking at genes less up regulated genes (Table. 4. 2). This trend is largely replicated in *swn-7* and *clf-28* (Table. 4. 2).

Interestingly, genes found most highly down regulated (top 10%) in *swn-7* and *clf-28* mutants showed over twice the genome average of percentage of genes possessing H3K27me^3 47% and 45% respectively (Table. 4. 1). The top 10% of genes down regulated in *swn-7 clf-28* double mutants showed only a negligible increase of genes possessing H3K27me^3 (20%) compared to the genome average (17%) (Table. 4. 1). Could the Pc-G act to promote gene expression? This could very well be the case, and the data presented here could be the first documented evidence that the Pc-G complexes could indeed used the H3K27me^3 mark as signal for gene activation in euchromatin. Considering the analysis of up regulated genes indicated that direct Pc-G targets were predominantly the most highly mis-expressed in *swn-7, clf-28,* and *swn-7 clf-28.* This suggests that *CLF* and *SWN* acting independently appear to be capable of both gene

activation and repression though histone methylation, whereas *SWN* and *CLF* acting redundantly appear to act preferentially to repress gene expression. These results indicate that the top 10% most up regulated genes were much more likely to be direct Pc-G targets, of *SWN CLF* acting redundantly, than the total of genes showing two fold increase in expression and possessing the H3K27me^3 mark. The over expression of *AG* has been shown to be the major cause for the *clf-* phenotype (Chanvivattana *et al.,* 2004; Goodrich *et al.,* 1997), indicating that the phenotypes of the Pc-G mutants is ectopic/up regulated gene expression so this is of more interest, but the Pc-G maybe capable of activating gene expression.

Developmental processes were of major interest considering the morphological defects observed in *swn-7, clf-28,* and *swn-7 clf-28* mutants are caused by gene mis-expression. Up regulated genes found in *swn-7, clf-28,* and *swn-7 clf-28* that possessed H3K27me^3 were analyzed to discover their biological significance, as known targets were found up regulated. In addition, the phenotypes of *swn-, clf-* and *swn- clf-* are most likely due to the over-expression of target genes.

Overall, the up regulated genes possessing H3K27me^3 in *clf-28* mutants showed an over-representation of genes involved in *ros* metabolism, trichoblast differentiation, response to biotic stimuli, lignin biogenesis, and most importantly flower development pathways. The top 10% of genes that are up regulated and possessing H3K27me^3 in *clf-28* shows an over-representation of genes involved in flower development (Table. 4. 3). This is consistent with the known role of *CLF* in repressing the floral homeotic genes *AG*, and *AP1* (Schubert *et al.,* 2006; Chanvivattana *et al.,* 2004; Goodrich *et al.,* 1997).

Total genes up regulated possessing H3K27me^3 in *swn-7* show over-representation in cell redox homeostasis and electron transport processes. The partitioning of genes possessing H3K27me^3 that are up regulated show that there is no over-representation of any particular biological processes (Table. 4. 3).

The up regulated genes possessing H3K27me^3 in *swn-7 clf-28* mutants show an over-representation in many biological processes including transcription, embryonic development, nucleic acid metabolism, response to stimulus, lipid transport, response to hormone stimulus, organogenesis, lipid metabolism, post-embryonic development, flower development, and lignin biosynthesis pathways. These processes are likely to represent the most influential aspects of development that are of Pc-G regulated. The top 5% up regulated genes found in *swn-7 clf-28* that possess H3K27me^3 are over-represented in embryo development, organogenesis, and lipid mobilization (Table. 4. 3).

The degree of up regulated expression of genes involved in developmental regulation and general transcription factors appears to reflect the morphology of the mutants analyzed. *swn*-7 shows no over representation of developmental processes and the *swn*-7 plants lack severe morphological defects. Whereas, *clf-28* seedlings show over-representation of genes involved in floral development, consistent with the *clf*-phenotype. The over representation of genes involved in embryo development in *swn-7 clf-28* seedlings therefore may reflect the direct cause of the *swn-7 clf-28* phenotype, and the up regulation of transcription factors that regulate this aspect of development. The large number of genes mis-regulated reflects the severity and complexity of the *swn-7 clf-28* phenotype. The mis-expression data for *swn-7 clf-28* shows a supreme number of genes encoding virtually every type of protein in the genome, but several patterns emerge that indicate the roles of the Pc-G in development that are discussed in more detail below.

4. 5. *SWN* and *CLF* act redundantly to repress the late embryogenesis transcriptional profile

swn-7 clf-28 mutants show increased expression of late embryogenesis traits (Chapter 3) indicating that Pc-G is responsible for repressing embryonic traits after germination in vegetative tissues. Both the master regulators, like *FUS3, LEC2, LIL, ABI3,* and *ABI4* and the down stream genes, e.g. *LATE EMBRYOGENESIS ABUNDANT1* (*EM1*), *ARABIDOPSIS THALIANA SEED GENE1* (*ATS1*), and *CRUCIFERINA* (*CRA1*) are

found up regulated in *swn-7 clf-28* (appendix: Table. 8. 2) This is in fitting with the finding with other points of evidence: firstly, the occurrence of the somatic embryos on the *swn- clf-* callus-like material (Chanvivattana *et al.*, 2004). Secondly, lipids specific to embryogenesis are found in *swn- clf-* callus-material (Chapter 3). Thirdly, the phenotypic and mis-expression similarity of *swn- clf-* to *pkl-* (Dean Rider S Jr et al., 2003), and *hsl2- hsi1-* (Tsukagoshi *et al.*, 2007), and the over-expression of *LEC1* (Stone *et al.*, 2008). Each of these mutants show de-repression of embryo specific traits in early seedling development. *pkl-* has recently been shown to possess reduced H3K27me^3 at the genes encoding seed storage proteins (Zhang *et al.*, 2008), strongly implicating that the *swn- clf-* and *pkl-* "pickled root" phenotype is a consequence of ectopic embryonic traits expression, and that *PKL* and the Pc-G function to repress their expression through H3K27me^3 (Zhang *et al.*, 2008). This is further supported by the embryonic trait master regulator *FUS3* has been shown to be regulated by the Pc-G as it is up regulated in *swn- clf-* and *mea-* tissues and *CLF* and *MEA* are found enriched at the *FUS3* loci (Makarevich *et al.*, 2006). The role of the Pc-G to mediate the expression of the whole embryonic maturation programme through H3K27me^3 mark isn't entirely novel, but this data confirms and extends this, by showing the scale of regulation of both master regulators of embryonic traits and down stream genes such as *LEC1*, LEAs and *EM1*. This data also illustrates the Pc-G is likely to directly regulate the master regulators and the down stream targets, which was previously unknown. The idea that the Pc-G regulates pathways is also illustrated in the regulation of flowering time and SAM maintenance.

4. 6. *SWN* and *CLF* act redundantly to regulate the SAM and lateral organs

There is a range of meristem specific genes that are required for SAM maintenance and differentiation state, for example, *STM, AIL5, AIL7, KNOTTED-LIKE1* (*KNAT1*), *KNAT2, KNAT6,* and *WUSCHEL-RELATED5* (*WOX5*) that are up regulated in *swn-7 clf-28* mutants (appendix: Table. 8. 2). In addition, genes that define the boundaries of the meristem, like *LBD4, LBD16, LBD40, LBD42*, and *CUC1* are found highly up regulated. This indicates that the Pc-G regulates the SAM region in multiple pathways.

There is no up regulation of known genes involved in leaf differentiation, and in fact, a *YABBY* family member (At2g26580) is down regulated, perhaps as a consequence of the lack of leaf tissue found in *swn- clf-* mutants. As these two classes of genes act antagonistically the increased expression of the meristem genes would reduce the expression of the leaf differentiators (Barkoulas *et al.*, 2007). It indicates the Pc-G represses genes required for meristem maintenance outside of the meristematic zones, and thus maintains both meristematic and differentiated tissue identity. However, without the knowledge of their detailed expression patterns in *swn- clf-* it is difficult to draw any conclusions.

4. 7. Analysis of gene mis-expression in *clf-28* with respect to its phenotype

Analysis of total genes up regulated in *clf-28* revealed that there were classes of genes that are likely to influence the *clf-* phenotype. Many genes encoding cell wall modifiers, e.g. *XYLOGLUCAN TRANSFERASE* (*XTR*) family member (At5g57530), and *EXPANSIN3* (*EXPA3*), however, many of the same classes of genes are down regulated in *clf-28* and these genes may be the driving force of the previously observed small cell size in roots of *clf-* plants (Kim *et al.*, 1998). These include *EXPA 9, EXPA14, EXPA18*, pectate lyase family proteins (e.g. At2g45220 and At5g04960), and *XTR9*.

Genes that respond to phyto-hormones are up regulated in *clf-28* e.g. *IAA31, AUXIN-INDUCED IN ROOT CULTURES 1* (*AIR1*), *MEDIATOR OF ABA-REGULATED DORMANCY 1* (*MARD1*), *GA REQUIRING 4* (*GA4*) and *RGA-LIKE 2* (*RGL2*) (appendix: Table. 8. 2). These genes are also likely to have an impact on the *clf-* phenotype, but the mechanism of how this may happen is unclear at the present.

A wide variety of genes are up regulated including transcription factors controlling floral organ identity, flowering time, meristem boundaries, and trichome morphogenesis, like: *AG, AP1, AP3, FLC, AGL19, MADS AFFECTING FLOWERING4* (*MAF4*), *SHATTERPROOF1* (*SHP1*), *SHP 2, SEPALLATA2* (*SEP2*), *SEP3, SEP4, SQUAMOSA PROMTER BINDING-LIKE PROTEIN3* (*SPL3*), *SPL4, FT, GLABRA2* (*GL2*), *C3HC4-*

and *GATA- TYPE* zinc finger family proteins, *MYB37, LBD25, NAC DOMAIN CONTAINING PROTEIN2* (*NAC2*), *NAC3, NAC55, NAC92, HOMEOBOX PROTEIN2* (*HB2)* and *HB40* (appendix: Table. 8. 2). These were the most predominant feature of up regulated genes found in *clf-28*, and are predominantly found in the top 10% of gene up regulated. The up regulation and probable ectopic expression of *AG*, which has been previously shown to cause the leaf curling of *clf-* plants (Goodrich *et al.*, 1997). *SEP2, SEP3*, and *SEP4* may also contribute to the leaf curling phenotype as the SEP proteins act in complexes with the AG proteins and required for *AG* activity (Favaro *et al.*, 2003). The increase in *FT, SPL3* and *SPL4* are likely to cause the early flowering phenotype of *clf-* mutants (Gandikota *et al.*, 2007; Schwarz *et al.*, 2008; Kardailsky *et al.*, 1999). Interestingly, there are flowering repressors, like *FLC, MAF4*, and *AGL19*, simultaneous up regulated in *clf-28*. This directly conflicts with the early flowering phenotype of *clf-* mutants, and the up regulation of flowering inducers, but illustrates how "linear" genetic pathways can be differentially regulated at several stages in Pc-G mutants. This point is further expanded in section 4. 9.

4. 8. The Pc-G regulates floral identity and MADs box gene expression

It becomes evident that a significant proportion of the MADs-box transcription factor gene clade appear to be regulated by *CLF* including *AG, AP1, AP3, FLC, AGL19, MAF4, SHP1, 2, SEP2, SEP3, SEP4* are up regulated in *clf-28* (appendix: Table. 8. 2). It was noted that there is a further regulation of the MADS-box gene clade with varied and unknown functions are up regulated in *swn-7 clf-28* including *AP1, PI, AGL5, AGL8, AGL9, AGL11, AGL14, AGL17, AGL42, AGL71,* and *AGL72* (appendix: Table. 8. 2). This may suggest a general regulation of the MADS-box clade by the Pc-G. However, some MADs-box genes are down regulated in *swn-7 clf-28* like *AGL12, AGL16*, and *AGL20*. Their down regulation is likely due to secondary effects or the Pc-G is required to positively regulate them directly or indirectly. This indicates that *CLF* predominantly regulates MADs-box genes but *SWN* acting redundantly with *CLF* also regulate regulates further MADs-box genes, suggesting a general role for the Pc-G in MADs-box gene regulation.

4. 9. The Pc-G regulates targets that act antagonistically in controlling flowering time

The role of the Pc-G in repressing flowering and floral organ identity outside of floral tissues is known (Goodrich *et al.*, 1997; Chanvivattana *et al.*, 2004) and is evident in the *clf-28* data which shows de-repression of flowering activators like *FT* and *SPL3*, however flowering repressors are also up regulated such as *FLC*. *swn-7 clf-28* mutants showed flowering repressors like *FLC* and *MAF4* are up regulated, and floral promoters are down regulated e.g. *FT*, *SPL3*, *SPL4* and *CONSTANTS* (*CO*) (appendix: Table. 8. 2). However, other floral promoters are up regulated like the enzymes that synthesis GA (*GA2-OX2*) (appendix: Table. 8. 2). GA does play many other roles in development so may or may not be specifically related to flowering, particularly in this context.

This illustrates that the Pc-G may regulate the flowering time pathway at several different steps, and that Pc-G targets may act antagonistically to one another.

The degree of Pc-G activity appears to affect flowering time, as knocking out *CLF* or *EMF2* results in differing degrees of early flowering suggesting that these genes are required to repress flowering. *EMF2* is more important in this process as it flowers earlier (Chen *et al.*, 1997) indicating that there is a difference in Pc-G activity. This suggests that the Pc-G is generally associated with repressing flowering in the absence of vernalization. However, it was observed that depleting Pc-G function further results in late flowering, as seen in *vrn2- emf2-10* double mutants which are late flowering, or fail to flower entirely (personal communication Dr. D. Schubert). The expression profile of *swn-7 clf-28* shows genes required for promoting flowering are reduced. This further suggests that the Pc-G also plays in integral role in the promotion of flowering.

This indicates that the different Pc-G complexes are required to co-ordinate both promotion and repression of flowering time. The Pc-G mutants suggest a gradient of Pc-G activity in which repressing flowering repressors is central to Pc-G function and is regulated by the redundant function of *SWN CLF* and *VRN2 EMF2*. Wherein, *SWN* or

VRN2 can maintain the central function in the *clf-* or *emf2-* mutants respectively. Considering the proposed role for the Pc-G in repressing and promoting flowering, the interaction of *FLC* and the Pc-G was looked at in more detail.

4. 9. 1. The Pc-G represses *FLC* independently of vernalization

The Pc-G is known to repress *FLC* following vernalization, in vernalization requiring backgrounds (Amasino *et al.*, 2005; Sheldon *et al.*, 2000). These mostly carry a *FRI+* allele to promote *FLC* expression (Johanson *et al.*, 2000). *FRI+* are very late flowering unless vernalization induces the Pc-G to repress *FLC* by the addition of the H3K27me^3 mark (Amasino *et al.*, 2005; Sheldon *et al.*, 2000; Gendall *et al.*, 2001; Wood *et al.*, 2006). Most laboratory strains are *fri-* and consequently have low *FLC* expression, subsequently they are early flowering and do not require vernalization to reduce *FLC* expression to permit flowering (Johanson *et al.*, 2000). Surprisingly, ChIP-chip studies found that H3k27me^3 mark is present at the *FLC* locus in both Ws and Col ecotypes (Fig. 4. 8) (Turck *et al.*, 2007; Zhang *et al.*, 2007), which are *fri-* early flowering laboratory strains. This raised the question of whether *fri-* lines are early flowering because the Pc-G represses *FLC* through histone modification in the absence of strong *FLC* promotion. However, the Pc-G mutants (e.g. *clf-* and *emf2-*) are early flowering, whereas late flowering would be expected if *FLC* is de-repressed in Pc-G mutants e.g. *clf-*. Therefore, there was a need to look at *FLC* expression in Pc-G mutants. *FLC* was found highly up regulated in the microarray of *clf-28* and *swn-7 clf-28*. To test this *FLC* expression was analyzed in independent biological samples of *clf-28, clf-81,* and *swn-7 clf-28* in Col *(fri-)* background (Fig. 4.2, 4.3, 4. 9 and Fig. 4. 10). This found that *FLC* expression is elevated in these Pc-G mutants. Why then are *clf-* plants early flowering with the increased expression of a potent floral repressor? It is presumably due to the ectopic expression of *AG* and *FT*, which are highly up regulated in the microarray data for *clf-28*. To further test the role of *FLC* in *clf-* plants the *clf-28 flc-3* double mutant was generated. *flc-3* mutants in a *Col* background are early flowering (Michaels & Amasino, 1999), much like *clf-*28, but the *clf- flc-* double mutants shows an enhanced *clf-* phenotype; It is earlier flowering (Fig. 4. 9) with increased leaf curling (Fig. 4. 10). In

long days the flowering defect was observable but mild in comparison to the short day growth conditions (Fig. 4. 9 A), where wild-type plants flowered with ~73 leaves, *flc*- with ~65 leaves, and *clf-28* with ~22 (Fig. 4. 9 B). The *clf-28 flc-3* mutants flowered with only ~9 leaves (Fig. 4. 9 B).

The leaf curling in *clf*- have been attributed to ectopic *AG* expression in the leaves (Goodrich *et al.*, 1997), the increased leaf curling in *clf- flc*- leaves indicates that *FLC* may act to repress *AG* expression in *clf*- plants.

This data provides evidence for; firstly, the *clf*- early flowering phenotype is suppressed by the over-expression of *FLC.* Secondly, the promotion of flowering in *clf*- plants occurs independently of the level of *FLC* expression. Thirdly, *CLF* is responsible for repressing *FLC* in wild-type plants. It demonstrates that the Pc-G regulates *FLC* independently of vernalization, and that the Pc-G may repress flowering time through *FLC* in wild-type plants perhaps contributing to the early flowering of *fri*- ecotypes. This illustrates neatly the Pc-G plays antagonistic roles in the flowering pathway; by repressing the flowing repressor *FLC* and other flowering promoters like *FT*. This demonstrates the duality of the CLF-Pc-G acting to promote and repress flowering.

4. 10. Potential causes of the delayed J-A transition and increased leaf initiation rate in *swn*- mutants

To discover a possible cause of the J-A and leaf initiation rate defects in *swn*- plants, the nature of up regulated genes in *swn-7* were analyzed, as *SWN* is a known repressor of gene expression (Chanvivattana *et al*., 2004; Wang *et al.*, 2006). The likely cause of the phenotype would be due to ectopic or over expression of target genes.

There are two cytochromes up regulated in *swn-7*, *CYP78A7* and *CYP82C4* (appendix: Table. 8. 2). Cytochromes are known to regulate the metabolism of the main phyto-hormones and secondary metabolites (Bishop *et al.*, 2006; Glawischnig, 2006; Hamberger & Bohlmann, 2006; Salchert *et al*., 1998). The *CYP78A7* mutants are wild-

type in phenotype (Wang *et al.*, 2008), but the over-expression of *CYP78A7* results in autonomous seed set and defects in ovule shape (Wang *et al.*, 2008). *CYP78A7* is closely related to *CYP78A5*, which is implicated in leaf initiation rate and J-A transition as it is a homologue of the *PLASTOCHRON1* (*PLA1*) gene in rice (Wang *et al.*, 2008; Miyoshi *et al.*, 2004). *pla1-* plants show increased leaf initiation rate and delayed J-A transition (Miyoshi *et al.*, 2004; Ahn *et al.*, 2002). *CYP78A5* is known to affect the J-A transition and leaf initiation rate as the *cyp78A5-* mutants exhibit accelerated leaf imitation rate (Wang *et al.*, 2008). Whereas, over-expressing of the genomic wild-type copy results in reduced leaf initiation rate (Helliwell *et al.*, 2001). However, the *cyp78a5- cyp78a7-* double mutants are early embryo lethal, indicating that *CPY78A5* and *CYP78A7* act redundantly. It is unlikely that the increased expression of *CYP78A7* in *swn-* could cause the delay in J-A transition or increased leaf initiation rate; in fact, the opposite effect may be expected as the increase in *CYP78A5* expression increase the leaf initiation rate (Wang *et al.*, 2008; Helliwell *et al.*, 2001). The increased fruit size and shape defects observed in the over expression of *CYP78A7* are not observed in the *swn-7* mutant, perhaps due to the increase in expression *CYP78A7* in *swn-* mutants is insufficient to cause the defects, and it is unknown if *CYP78A7* is mis-expressed in the fruits of *swn-* plants or just the seedling stage.

The up regulation of the cytochromes may not reveal the whole story, given that the up regulation may not be the in the apex and new leaves, the regions where the regulation of leaf initiation rate and the J-A transition occurs (Wang *et al.*, 2008). In addition, it is likely that other factors such as secondary metabolism, which are also found up regulated in *swn- 7* may effect the J-A transition and leaf initiation rate, either independent of cytochromes, transcription factors, hormonal changes, or interacting with them (Helliwell *et al.*, 2001; Wang *et al.*, 2008; Fahlgren *et al.*, 2006; Telfer *et al.*, 1997).

The only known pathways controlling the J-A transition are through *miRNAs* (*ta-siRNA*) acting to regulate *ARF3, ARF4, SPL3, SPL5, SPL9, SPL15* (Hunter *et al.*, 2006; Schwarz *et al.*, 2008) or through *CYP78* genes (Wang *et al.*, 2008), which in turn regulate the J-A

transition and the plastochron. None of these genes are mis-expressed in *swn-7* (appendix: Table. 8. 2), either *SWN* does not act to regulate these pathways directly, or, *SWN* acts further down stream, which may or may not be part of the same pathway, indicating that *swn-* may present a novel pathway regulating the J-A transition and leaf initiation. Alternatively, the age of the samples used for the microarray did not capture the key difference between wild-type and *swn-7* that cause the phenotypic defects. The *swn-* phenotypes are possibly a consequence of changes in metabolism, transcription factors, expression, or changes in hormone responses, which effect growth, J-A transition and leaf initiation.

4. 11. Analysis of gene mis-expression in *swn-7* and *clf-28* mutants, with respect to common *CLF* and *SWN* roles

Genes up regulated in *swn-7* and *clf-28* showed no over represented biological functions. This shows us that the common function of *SWN* and *CLF* appear not to repress pathways or processes, but rather individual genes in a variety of pathways and functions.

Genes down regulated in *swn-7* and *clf-28* showed over representation of biological processes including: lipid transport and storage (e.g. protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (At5g54740) and *OLEOSIN 2*), and response to endogenous stimulus (e.g. *SQUALENE EPOXIDASE 1, 1* (*SQP1, 1*), and zinc finger (GATA type) family protein (At4g26150) (appendix: Table. 8. 2).

4. 12. Gene mis-expression in *msi1-* and *emf2-* compared to *clf-* and/or *swn- clf-* to identify probable direct targets of the Pc-G

As *SWN* and *CLF* are part of the Pc-G complex it would be expected that the other components of the complex when mutated would result in the mis-expression of the same targets, and these similarly mis-expressed genes would represent true targets of the Pc-G, particularly those decorated with H3K27me^3. To find these true potential targets

of the Pc-G the comparison of the *msi1*- and *emf2*- expression to *clf-28* and/or *swn-7 clf-28* was carried out.

MSI1 is a member of the Pc-G and CAF-1 complexes (Kohler *et al.*, 2003a). Gene expression analysis showed a range of genes that are mis-expressed in co-suppressed *MSI1* seedlings (Hennig *et al.*, 2003), and the limited collection of genes published was compared to genes with H3k27me^3 as this mark is associated with Pc-G function. There were 135 mis-expressed genes in co-suppressed *MSI1* plants of which, 106 genes were reported as being up regulated. 44.3% (47) of the 106 genes possessed H3k27me^3. 37/106 genes up regulated in co-suppressed *MSI1* plants and up regulated in *swn-7 clf-28* and 51.4% (19) of these possessed H3K27me^3 (Table. 4. 4 and appendix: Table. 8. 3). Genes of most interest include *PLANT DEFENSIN 1.2B* (*PDF1.2b*), *IDENTICAL TO LATE-EMBRYOGENESIS ABUNDANT M17 PROTEIN* (At2g41260) *CRA1*, a heat shock protein-related (At5g47600), a *LTP* family protein (At5g55560), *LTP3*, a zinc finger (C3HC4-type ring finger) family protein (At5g60250), and a 12s seed storage protein (appendix: Table. 8. 3).

The percentage of genes possessing H3K27me^3 is higher in the genes found up regulated in *msi1*- and *swn-7 clf-28* than in the genes in co-suppressed *MSI1* consistent with the know role of *MSI1* as a member of the Pc-G. The genes up regulated in both *msi1*- and *swn-7 clf28* are very likely direct Pc-G targets.

emf2-1 array data, as carried out and kindly donated by Dr. R. Sung (data in part published (Moon *et al.*, 2003a) (and analyzed in the same manner as described (see materials and methods)) was compared to *clf-28*, and *swn-7 clf-28* arrays. These data sets were compared because *EMF2* is a core member of the Pc-G post-germination and the *emf2*- phenotype is in many respects similar to *clf*- (Chanvivattana *et al.*, 2004), only more severe. The *emf2-1* array data shows that 1334 genes were up regulated after 15 days growth and this number appears to be an intermediate between *clf-28* and *swn-7 clf-28*, which may reflect the intermediate phenotype of *emf2-1* compared to the "weaker" *clf-28* and the "stronger" *swn-7 clf-28* phenotypes. Perhaps, it indicates a

gradient of Pc-G function with respect to developmental regulation. It may also indicate that there is more redundancy of *SWN* and *CLF* than between *VRN2* and *EMF2,* as *EMF2* appears to play a more indispensable role in the Pc-G than *CLF*. However, the increased gene mis-expression in *emf2-1* may be the consequence of the extra days of growth resulting in the increased ectopic expression of secondary effects. This point is supported by the percentage of genes that possess H3K27me^3 is slightly higher than the genomic average, but notably less than genes up regulated in *swn-7, clf-28,* or *swn-7 clf-28* genotypes possessing H3K27me^3 (Table. 4. 4). The comparison of genes up regulated in both *emf2-1* and *clf-28* seedlings revealed that 77 genes were up in both genotypes and. These 77 genes are very likely direct and indirect targets of *EMF2* and *CLF* (Table. 4. 4). 34 (44.2 %) of these 77 genes possessed H3K27me^3 (appendix: Table. 8. 3), considerably higher than the genome average of 17%. The intersection of 34 genes up regulated in *clf-28* and *emf2-1* that possess methylation include 6 MADs domain proteins, e.g. *AG, AP3, AGL1, AGL5, AGL9, FLC, AG,* and *AP3* are required for flowering and carpel development, and are known to be both targets of *CLF* and *EMF2* (Chanvivattana *et al.*, 2004) (appendix: Table. 8. 3). This supports the hypothesis that the Pc-G regulates the *AGL* clade. Other genes of interest included *FT, BETA-AMYLASE1,* a pectinesterase family member (At2g45220), and a LTP (At4g12550) (appendix: Table. 8. 3). This supports the role of the Pc-G in lipid mobilization, cell wall expansion, sugar metabolism, and flowering.

378 genes were found up regulated in *swn-7 clf-28* and *emf2-1* and of these 23% (89) of the genes possessed H3K27me^3 (Table. 4. 4 and appendix: Table. 8. 3). This is marginally above the genome average, but relatively low, and indicates that like *swn-clf-* the majority of gene mis-expression is due to secondary effects, perhaps due to the de-repression of master regulators of transcription programmes. Genes mis-expressed include a collection of MADs box, homeodomain genes, metabolic enzymes like *AGL1, AGL9, AGL11, SEP2, PI, FLC, AG, AP3, AP1, LTP3, BELLRINGER, CRC, STM, KAN2, FLS1, LRP1,* and *BMY1* (appendix: Table. 8. 3). Of the genes up regulated in both *emf2-1* and *swn-7 clf-28* and possess H3K27me^3 20% (18 genes) are in the top 20%

of genes up regulated in the *swn-7 clf-28* mutant suggesting that they are true Pc-G targets (appendix: Table. 8. 2), supporting the previous finding that direct Pc-G targets are the predominantly most highly up regulated genes.

The genes found up regulated in *emf2-1* and *clf-28* and/or *swn-7 clf-28* possessing H3K27me^3 are highly probable as direct targets of the Pc-G. This suggests that the increased number of genes up regulated are secondary effects of the ectopic expression of direct targets, and the true targets are more probably found in the intersection of the *emf2-1* and *clf-28,* and *emf2-1* and *swn-7 clf-28.* The fact that more up regulated genes of *emf2-1* are found in *swn-7 clf-28* than *clf-28* explains why the *emf2-1* phenotype is more severe than *clf-28,* this also demonstrates that *EMF2* acting discretely may play a broader role in development than *CLF* acting discretely, in terms of regulating flowering and morphological development. This data in part confirms the potential role for the Pc-G not only to be involved in regulating flowering time and floral organ identity genes but also cell wall modification, and aspects of metabolism.

4. 13. Discovering causative agents of the *swn- clf-* phenotype

The *hsi2- hsl1-* double mutant causes the ectopic expression of embryonic traits and stunted growth, phenotypic attributes reminiscent of the *swn- clf-* (Tsukagoshi *et al.,* 2007; Chanvivattana *et al.,* 2004). The comparison of the *hsi2- hsl1-* and the *swn-7 clf-28* mis-expression data could identify a selection of candidate genes that may be responsible for aspects of the *swn- clf-*mutant phenotype.

There are 823 genes up regulated in the *hsi2- hsl1-* double mutant, 35.6% (293 genes) possess the H3K27me^3 mark (appendix: Table. 8. 3), over twice the genome average (Table. 4. 4). *HSI2 HSL1* are not mis-expressed in Pc-G mutants, and therefore are unlikely to be direct or secondary Pc-G targets. *HSI2 HSL1* appears to play a similar role in repressing embryonic traits as the Pc-G. The comparison of genes up regulated in *hsi2- hsl1-* and *swn-7 clf-28* revealed that over half of the genes up regulated in *hsi2- hsl1-* are up regulated in *swn-7 clf-28* (471 genes) and 38% (179 genes) possess

H3K27me^3 (Table. 4. 4 and appendix: Table. 8. 4). This feature may be relatively unsurprising when considering which genes are similarly up regulated; they include some of the master regulators of embryonic traits like *FUS3, ABI3,* and *ABI4,* and down stream genes like LEAs and *EM1* (appendix: Table. 8. 4). It illustrates that the phenotype of *swn-7 clf-28* is likely dependent on the ectopic expression of embryonic traits, at least in part caused by these genes, especially as *FUS3, ABI3,* and *ABI4* are in the top 10% up regulated in *swn-7 clf-28*.

The similarity of data sets may indicate two obvious points: 1) The *HSI2* and *HSL1* transcription factors are likely to function, in some way, in conjunction with the Pc-G perhaps guiding the Pc-G to the seed maturation targets post-germination. (2) The Pc-G modulates aspects of sugar metabolism to repress sugar from inducing the seed maturation program as the *hsi2- hsl1-* mutant phenotype is only revealed when grown on high sugar concentrations, and the *swn-2 clf-81* phenotype is enhanced in when grown with high sugar concentration (Chapter 3).

4. 14. Is there a function equivalent of the PRC1 complex in *Arabidopsis*?

In flies, the PRC2 complex carries out only one stage to bring about stable repression of target genes. It is believed that the PRC2 complex recognizes the PRE and catalyzes the H3K27me^3 mark. This mark is then "interpreted" by the PRC1 which then creates a "closed" region of chromatin which is transcriptionally silent (Schmitt *et al.*, 2005). This mechanism has not been fully conserved in plants, as no homologous components of the PRC1 are found in the *Arabidopsis* genome. The H3K27me^3 mark is present in plants as is the PRC2 complex, and if no homologous PRC1 complex is present then perhaps a functional equivalent is acting to maintain the stable repression of the PRC2 targets in plants. There are several possibilities of how this occurs, the most likely system being chromatin remodellers, as several have been shown to actively repress gene expression. Likely candidates were selected on the based on similarity of their mutant phenotypes relative to Pc-G mutant phenotypes, they and include *EMF1,* the CHD3 type chromatin remodeller *PKL,* and *TFL2/LHP1*. This hypothesis was tested by comparing the mis-

expression (or proposed direct targets) of chromatin remodelling complexes and Pc-G mutants, because if these carry out a PRC1-like role then the mis-expression of the same genes would be expected to be up regulated in PRC2 mutants as in the PRC1 mutants.

Genes found mis-expressed in *emf1-* alleles were compared to *clf-28, swn-7 clf-28* expression array and the H3K27me^3 epigenome map because *EMF1* is involved in Pc-G function probably in the same genetic pathway as *EMF2* (Calonje *et al.*, 2008).

The analysis of data revealed that >1000 genes were at least 2 fold up regulated in the two *emf1-* mutant alleles, *emf1-1* is phenotypically weaker than *emf1-2* (Moon *et al.*, 2003a). However, the true direct targets should be up regulated in both alleles of *emf1-* and the percentage of genes up regulated with histone methylation was much higher in the genes up regulated in both *emf1-1* and *emf1-2* than the alleles analyzed separately (data not shown). Therefore, only those genes up in *emf1-1* and *emf1-2* were compared to up regulated genes in *clf-28* and *swn-7 clf-28*.

288 genes were found up regulated in *emf1-1* and *emf1-2* and 20.1% (57) possessed H3K27me^3 (Table. 4. 4 and appendix Table. 8. 4), which is marginally above the genome average. The comparison of genes up regulated in the *emf1-* alleles and *clf-28* revealed only 14 genes. Despite the minimal overlap, the genes that are overlapping show a very high percentage of genes to be methylated (42.9% (6)) (Table. 4. 4 and appendix: Table. 8. 4). This gives the appearance that these genes are very likely to be Pc-G regulated. The 14 genes comprised genes lacking H3K27me^3 include *SPL4* and an LTP family member (At4g12510). The genes that possess H3K27me^3 and are up regulated in *clf-28* and *emf1-* alleles include nodulin MtN21 family protein (At2g39510), *AGL5*, asparaginyl endopeptidase (At3g20210), *BMY1*, and two expressed proteins (At1g16950, At2g42610) (appendix: Table. 8. 4.).

The comparison of genes at least 2 fold up regulated in *emf1-* alleles and *swn-7 clf-28* showed 140 genes were up regulated in each of the mutants (Table. 4. 4). 30/140 (21.4%) up regulated genes possesses H3K27me^3 methylation (Table. 4. 4 and appendix:

Table. 8. 4)), marginally above the genome average. These 30 genes are very likely targets and show very similar characteristics to the top 10% of genes up regulated in *swn-7 clf-28*. Genes found are involved in carpel and seed development and meristem maintenance are, for example, *CRA1*, 2S albumin storage proteins, *PI, CRC, AGL5, STM,* and *AIL5* (appendix: Table. 8. 4). In addition, an auxin synthase gene (*YUC4*) is also up regulated (appendix: Table. 8. 4). The similarity of these genes up regulated may shed light on the callus-like material formed in *swn-7 clf-28* mutants as it has been observed that the *emf1-2* mutants also produces callus projections in the aerial parts (Calonje *et al.*, 2008).

There are an incredibly small number of genes commonly up regulated in *clf-28* and *emf1-* alleles when considering the early flowering time and mis-regulation of flower identity, and *EMF1* is known to repress *AG* (Calonje *et al.*, 2008). Known targets of *CLF; AG, FLC,* and *AP1* are found to possess H3K27me^3 and are mis-expressed only in *emf1-2* allele. This is the phenotypically weaker *emf1-* mutant allele (Calonje *et al.*, 2008), suggesting that the *swn- clf-* phenotype is not due to the up regulation of discrete *CLF* targets. Instead, the increased severity in the *swn- clf-* double mutant is caused by a shift in the transcriptional profile, from the de-repression of flowering time and flower identity genes in *emf1-2, emf2-1,* and *clf-28* single mutants, and the *emf1-1* and *swn-7 clf-28* mutant caused the de-repression of embryonic traits and stem cell fate genes.

The *pkl-* mutant displays a “pickle” root phenotype root tip and the de-repression of embryonic traits (Ogas *et al.*, 1997; Henderson *et al.*, 2004; Li *et al.*, 2005) is superficially identical to *swn-7 clf-28* mutants (chapter 3) (Chanvivattana *et al.*, 2004), therefore, genes found commonly up regulated would also be potential causative agents of the similar phenotype. As *PKL* is a chromatin remodeller responsible for repressing gene state this suggested that *PKL* might be directly involved in repressing common target genes as Pc-G, particularly as the PKL protein possess a PHD domain which has been implicated in binding to methylated histone tails in mice (Papait *et al.*, 2008). To test this I compared microarray data from the *pkl-* plants whose expression was up regulated in a *PKL* dependent manner, but independent of the effects of GA (Li *et al.*,

2005). 19.4% (55 genes) of the 284 two fold up regulated genes in *pkl*- possessed H3K27me^3 (Table. 4. 4), which is marginally above the genome average. There are 64/284 genes that are up regulated in both *pkl*- and *swn-7 clf-28* and 40.6% (36 genes) of these genes possess H3K27me^3, this is >2 greater than the genome average (Table. 4. 4 and appendix: Table. 8. 3).

This suggests that chromatin modification and chromatin remodelling processes are, in some way, mechanistically connected. This becomes more transparent when looking at which genes are commonly up regulated, as 7/64 genes are directly related to embryogenesis storage proteins (appendix: Table. 8. 4). It appears that *SWN*, *CLF*, and *PKL* plays a major role in embryonic trait repression, i.e. *PKL* may be acting as a functional equivalent of the PRC1 in plants, with respect to embryonic traits, but whether they act in the same mechanistic pathway is unclear with this type of analysis. This hypothesis was recently shown, in *pkl*- mutants as the increase in the embryogenesis traits is found to correlate with a reduction in H3K27me^3, suggesting that *PKL* is required to maintain Pc-G mediated H3K27me^3 marks in the embryonic trait pathway (Zhang *et al.*, 2008).

The master regulators of embryogenic traits like, *FUS3* and *LEC2* are found up regulated but show a dependence on GA and so were discounted from comparison, as it is unlikely the Pc-G plays a definitive role in regulating the GA response. This is based on the observation that *swn- clf*-mutants show no phenotypic effect when grown on GA inhibitor or GA itself (chapter 2 and data not shown).

The TFL2/LHP1 protein was characterized to identify target genes on a whole chromosome 4, i.e. where TFL2 was found to be bound on chromosome 4. The loci bound by TFL2 was found to correlate with the H2K27 me^3, to potentially repress loci expression (Turck *et al.*, 2007). This was supported by *TFL2* being required for *FLC* repression post-vernalization (Mylne *et al.*, 2006). Given this evidence it was proposed that *TFL2* is potentially a functional equivalent of the PRC1 complex, acting to repress target genes of the PRC2 (Turck *et al.*, 2007). To test this hypothesis I carried out

comparison of genes that TFL2 was found bound to and genes up regulated in *swn-7 clf-28*. TFL2 was found to bind to 538 loci on chromosome 4, of these 427 genes (79%) possess H3K27me^3; the genome average of H3K27me^3 is 17%, which strongly implicates the co-localisation of H3K27me^3 and TFL2 binding. There are 65 genes enriched by TFL2 that are up regulated in *swn-7 clf-28* (Table. 4. 5). 87.7% (57) of these genes possess H3K27me^3 (Table. 4. 5 and appendix: Table. 8. 4), this is higher than the genome average and the percentage of H3K27me^3 possessing genes found to co-localized with TFL2. This could suggest that the Pc-G and *TFL2* do act at common targets perhaps the same mechanism to bring about the stable repression of genes found in euchromatin, but due to the high percentage of genes possessing H3K27me^3 bound by TFL2 is incredibly high, it is therefore, difficult to draw conclusions as to whether *TFL2* and the Pc-G mechanistically interact. This is especially true as only 57 genes of the 427 genes possessing H3K27me^3 and enriched by TFL2 are up regulated in *swn-7 clf-28*. This suggests that *TFL2* does bind to some of the same targets as the Pc-G, which the Pc-G acts to repress, but there are still a vast number of target loci that are bound by *TFL2* which are probably independent of the Pc-G. There are also many loci that are Pc-G targets that are not targets of *TFL2*. This suggests that *TFL2* is not the only functional equivalent of the PRC1, or not all of the PRC2 targets are targets by the PRC1 like complexes.

In conclusion, none of the chromatin remodellers analyzed here regulates direct Pc-G targets on a genome scale, but rather each seems to regulate a sub-section of Pc-G targets. This indicates that there may not a single functional PRC1 equivalent but multiple chromatin remodellers have been recruited to Pc-G targets to carry out the PRC1-like role. The genes that have been found to be regulated in the Pc-G mutants and chromatin remodellers may represent neat model loci to investigate the mechanism underlying their repression.

4. 15. Summary and conclusions

To gain insight into which biological processes are Pc-G regulated a genome wide gene expression analysis was carried-out on *swn-7, clf-28,* and *swn-7 clf-28* compared to wild-type seedlings. Ectopic gene expression in each of these mutants was analyzed to find which biological functions are being mis-expressed, and to discover potential direct targets and secondary targets. These derived gene lists were then compared to other relevant array data, in an attempt to derive insight into the mechanism of Pc-G gene repression.

Genes mis-regulated in the mutants that possessed H3K27me^3 were considered as likely targets for the Pc-G, as SWN and CLF create and maintain virtually all of H3K27me^3 in euchromatin. This revealed that *CLF* has a potential of 128 discrete targets, and *SWN* has a potential of 109 discrete targets. *CLF* and *SWN* have potential 91 targets in common. *CLF* and *SWN* have 1004 potential redundant targets. Genes lacking histone methylation but with altered expression are likely to be secondary targets. Although the tissue used only reflects a brief moment in the developmental cycle and histone modification is known to be a dynamic process, genes that have lost H3K27me^3 from an early developmental stage and loci that acquire H3K27me^3 later in development may not be observed. Direct targets were predominantly the most highly up regulated.

Using the array data and its comparison to the H3K27me^3 epigenome map I have identified 1333 novel Pc-G targets, and have validated 13 potential novel direct targets more than doubling the previously known Pc-G targets. *similar to PRE1* (At5g15160) was validated as being up regulated in independent *swn-2,* and *swn-7* alleles was considered a direct targets because it was found in the top 10% of genes up regulated *swn-7* and possess H3K27me^3. This gene is probably not the major cause of the *swn-* phenotype but may contribute to them. *LEC1, FUS3, ABI3, EM1, CUC2, LBD16, LBD40, AIL5, AIL7, GAOX, YUC4, AGL16, FT,* and *FLC* were validated as being up regulated in an independent *swn-7 clf-28* sample. These were selected as they were in the top 10% of genes up regulated in *swn-7 clf-28* array data set, and they possess the

H3K27me^3 mark, representing potential targets. These genes potentially cause aspects of the *swn- clf-* phenotype, through regulation of stem cells, lateral organ boundaries and the up regulation of embryonic traits.

The array data showed that the majority of gene mis-expression was due to secondary effects, and that the direct Pc-G targets were the most highly up regulated. The genes that were the highest up regulated in *swn-7, clf-28,* and *swn-7 clf-28* correlated directly with the highest percentage of H3K27me^3 possessing genes, overall. Those genes that are less highly up regulated show reduced percentage of genes possessing H3K27me^3. Many of the genes most up regulated in *swn-7 clf-28* were developmentally important genes, involved primarily in embryogenesis maturation, flowering, and differentiation.

clf- mutants are early flowering, indicating its role in repressing flowering but *FLC* was found up regulated in *clf-* alleles and *swn-7 clf-28* mutants. *FLC* is known to posses H3K27me^3 in non-vernalization requiring ecotypes (Zhang *et al.*, 2007; Turck *et al.*, 2007). This indicated that the Pc-G both plays roles in promoting flowering and repressing flowering independently of vernalization. To uncover the role of *FLC* in *clf-* mutant the *clf-28 flc-3* double mutant was generated. The *clf- flc*-mutants show earlier flowering than either parent, this shows that that *FLC* delays flowering in *clf-* plants, indicating that *clf-* plants flower independently of the level of *FLC,* probably through ectopic expression of *AG* and *FT*. It suggests that the CLF-Pc-G acts to repress *FLC* expression independently of vernalization in wild-type plants, and therefore probably acts to promote flowering. This demonstrates that the Pc-G acts in the flowering pathway but on antagonist genes.

To validate the discovery of novel targets, other mutant Pc-G members mis-expression were compared to *clf-28* and/or *swn-7 clf-28* mis-expression data. Genes that are up regulated in multiple Pc-G mutants and possess H3K27me^3 strongly represent potential direct targets of the Pc-G. Genes with roles in flowering, cell expansion, carbohydrate metabolism, and seed storage proteins were the most predominant feature of gene found commonly up regulated in *emf2*-1 and *msi1-* , and *clf-28* and *swn-7 clf-28* Although the

common genes up regulated in *swn-7 clf-28* and each of the other Pc-G mutants show little overlap with each other. This may illustrate that the Pc-G is a more dynamic complex than originally conceived, as each of the components display a functional overlap with the catalytically active components, *SWN* and *CLF*, but they also display independent gene expression. This may suggest that the core members acting in cohorts do not only dictate Pc-G gene regulation, but also the individual members display independent roles in development.

The phenotype of *swn-7 clf-28* is probably a consequence of mis-regulation of master regulators of embryogenesis, like *FUS3* and *ABI* genes as not only are these gene the top 10% of up regulated genes, but similar aspects of the *swn-7 clf-28* phenotype is observed in *pkl-* and *hsi2- hsl2-* mutants which show increased expression of these genes.

This strongly indicates that the strong *swn- clf-* mutant phenotype is correlated with ectopic expression of key embryonic regulators..

However, mis-expression of *FUS3* and *ABI* genes does not explain all aspects of the *swn- clf-* phenotype such as the callus proliferation, or the root hair deformation. These aspects may be a consequence of mis-regulated meristem genes, floral and root specific transcription factors e.g. *STM, AG, AIL5,* and *AIL7.* This suggests that *swn-7 clf-28* mutants are in a state of embryo, floral organ differentiation, and stem cells fate. These cell fates may be tissue specific, or overlapping. It remains unclear how the embryonic and stem cell expression networks interact, and this may explain the tissue specificity, or the callus-like appearance. It is impossible to draw solid conclusions based solely on the expression data shown here as the tissue specific gene mis-expression has been shown in *swn- clf-* mutants (Personal communication Dr. O. Clarenz). In conclusion, it is probable that the *swn- clf-* phenotype is due to tissue specific transcription profiles being ectopically expressed causing the disruption of previously established tissue specific transcription profiles. This illustrates the general developmental function of the Pc-G is to repress tissue specific transcription profiles in tissues that they should not be expressed.

Regulation of embryogenesis and stem cell maintenance, and aspects of flowering are likely to be regulated by chromatin remodellers (*PKL* and *TFL2*), and possibly the CAF-1 complex (*MSI1)*. The roles of chromatin remodellers appear to overlap partially in function with the Pc-G, but certainly not entirely. The idea of the these chromatin remodellers mechanistically interact with Pc-G to repress common target genes is supported by the fact that genes up regulated in mutants possess a marginally higher than expected percentage of genes possessing H3K27me^3, which in turns increases when analysing genes up regulated in *swn-7 clf-28* and chromatin remodeller mutants. The fact that each of the factors only partially function at likely Pc-G targets reveals that it maybe unlikely that there is a single mechanism to stably regulate PRC2 targets in a PRC1-like manner, at least not the complexes investigated here.

Now we have come to the point where all the data has been exhibited, so now is time to evaluate what has been uncovered, and discuss the possible implications of the data in the following chapter.

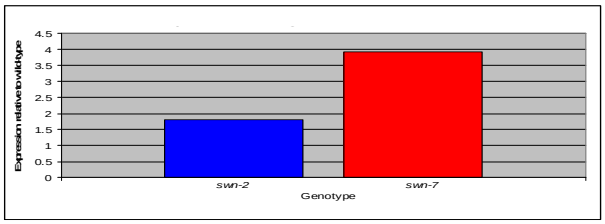

Fig. 4. 1. Q-PCR analysis of gene expression in independent *swn-* alleles.
RNA was isolated from whole seedlings of two independent alleles of *swn-2*, and *swn-7* at 10 days old. Two biological replicates were taken, the cDNA of these were used for Q-PCR analysis, Q-PCR reactions were carried out in triplicate. Results of the two biological replicates were averaged, and then normalized to *EIF-4* (a commonly accepted reference gene) this was then compared to wild-type replicates (which underwent the same normalization to *EIF-4*). Both *swn-* alleles show elevated expression levels of At5g15160 (*similar to PACLOBUTRAZOL RESISTANT1*).

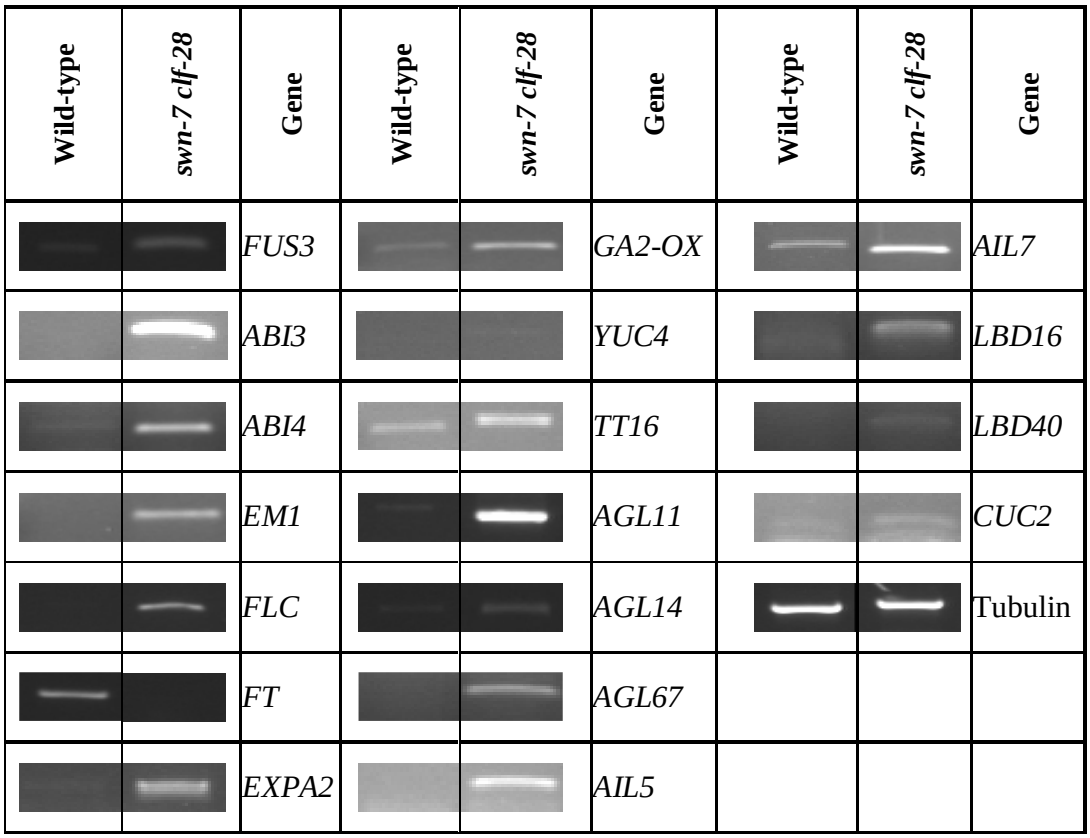

Fig. 4. 2. RT-PCR analysis of gene expression in *swn-7 clf-28*.
RNA was isolated from whole seedlings of wild-type at 10 days and *swn-7 clf-28* mutants at 12 days old (independent of the microarray samples) and tested by RT-PCR (see methods and methods) to observe the expression differences predicted in microarray data set. Each of the genes tested are found up regulated in *swn-7 clf-28* compared to wild-type, except *FT* which is down regulated, consistent with the microarray data.

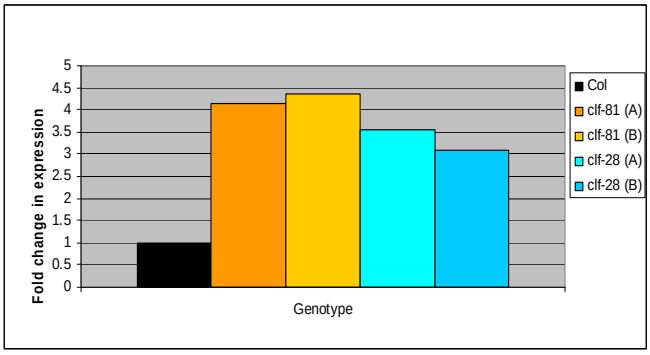

Fig. 4. 3. Q-PCR analysis of *FLC* expression in independent *clf-* alleles.
RNA was isolated from whole seedlings of two independent alleles of *clf-81*, and *clf-28* at 10 days old, Q-PCR analysis revealed both *clf-* alleles have elevated *FLC* expression compared to wild-type plants as normalized to *EIF-4*.

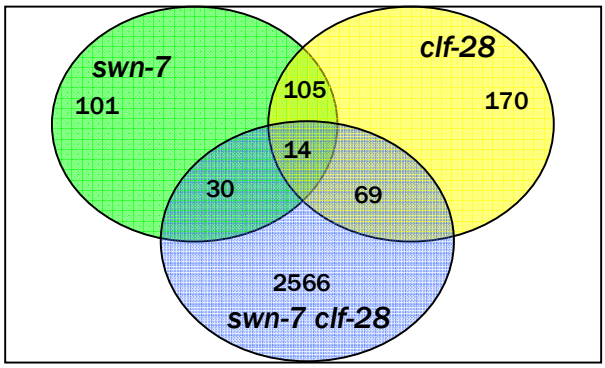

Fig. 4. 4. Comparison of up regulated genes in Pc-G mutants.
Venn-diagram showing the number genes that are up regulated (in with a minimum of two fold difference) in *swn-7*, *clf-28*, and *swn-7 clf-28* compared to wild-type. *swn-7* =252, *clf-28*= 360, and *swn-7 clf-28*= 2701 genes up-regulated.

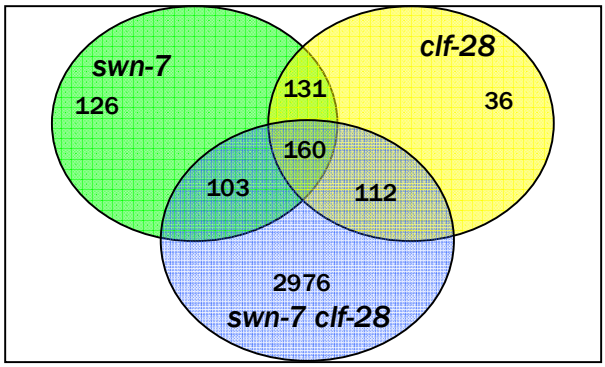

Fig. 4. 5. Comparison of down regulated genes in Pc-G mutants.
Venn-diagram showing the number of genes that are down regulated (with a minimum of two fold difference) in *swn-7*, *clf-28*, and *swn-7 clf-28*. *swn-7* =529, *clf-28*= 444 compared to wild-type, and *swn-7 clf-28*= 3376 genes down-regulated.

	Genotype		
	swn-7	***clf-28***	***swn-7 clf-28***
Total gene mis-expressed	781	804	6077
Genes total mis-expressed with H3K27me^3	190	216	1133
% genes with H3K27me3	**24.3**	**26.9**	**18.6**
Genes up regulated	252	360	2701
Genes up with H3K27me^3	55	112	557
% genes with H3K27me3	**21.8**	**31.1**	**20.6**
Genes down regulated	529	444	3376
Genes down with H3K27me^3	135	105	576
% genes with H3K27me3	**25.5**	**23.6**	**17.1**

Wild-type genome average.	
Total genes	27001
Genes with H3K27me^3	4596
% genes with H3K27me3	17.0

Table. 4. 1. Mis-expressed genes in *swn-7, clf-28,* and *swn-7 clf-28* that possess H3K27me^3.
Analysis of genes mis-expressed, up and down regulated compared to genes known to possess H3k27me^3 and the genome average (Zhang *et al.*, 2007).

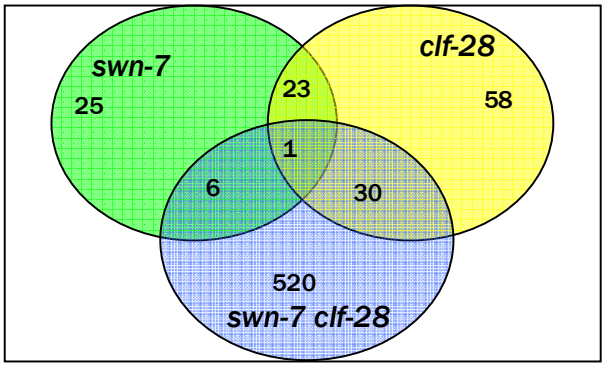

Fig. 4. 6. Comparison of genes up regulated in Pc-G mutants that also possess H3K27me^3 Venn-diagram showing the number of genes that are up regulated in *swn-7*, *clf-28*, and *swn-7 clf-28* that possess H3K27me^3. *swn-7 = 55*, *clf-28*= 112, and *swn-7 clf-28*= 557 genes up-regulated.

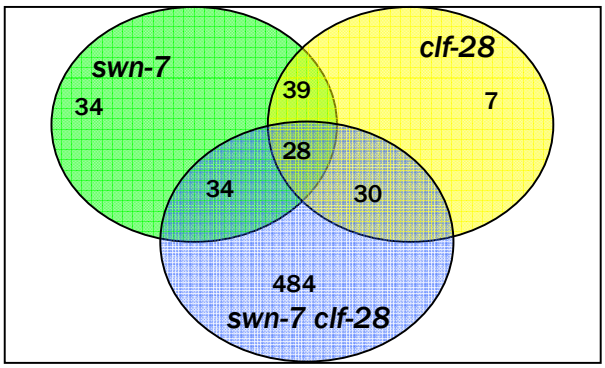

Fig. 4. 7. Comparison of genes down regulated in Pc-G mutants that also possess H3K27me^3.Venn-diagram showing the number of genes that are down regulated in *swn-7*, *clf-28*, and *swn-7 clf-28* which also possess H3K27me^3. *swn-7 =135*, *clf-28*= 104, and *swn-7 clf-28*= 576 genes down-regulated.

Genotype	% of up regulated genes	Total genes	Genes with H3K27me^3	% genes with H3K27me3
swn-7	5%	14	5	**35.7**
	10%	29	8	**27.6**
	10-20%	28	5	**17.9**
	20-50%	85	15	**17.6**
	50-100%	108	27	**25.0**
	Total	252	55	**21.8**
clf-28	5%	18	10	**55.6**
	10%	36	19	**52.8**
	10-20%	35	16	**45.7**
	20-50%	107	37	**34.6**
	50-100%	180	40	**22.2**
	Total	360	112	**31.1**
swn-7 clf-28	5%	138	75	**54.3**
	10%	274	156	**56.9**
	10-20%	208	93	**44.7**
	20-50%	847	180	**21.3**
	50-100%	1340	128	**9.6**
	Total	2701	557	**20.6**
Genotype	**% of down regulated genes**	**Total genes**	**Genes with H3K27me^3**	**% genes with H3K27me3**
swn-7	10%	53	25	**47.1**
	Total	529	135	**25.5**
clf-28	10%	44	18	**40.9**
	Total	444	105	**23.6**
swn-7 clf-28	10%	336	70	**20.8**
	Total	3376	576	**17.1**
Genome average	N/A	27001	4596	**17.0**

Table. 4. 2. Genes partitioned by level of mis-expression and percentage of genes with H3K27me^3 in *swn-7, clf-28,* and *swn-7 clf-28* mutants.

Genes found up regulated, more than two fold, in each mutant background were sorted according to their level of expression and partitioned into the top 5%, 10%, 10-20%, 20-50%, and 50-100% in terms of the total genes up regulated. These sections were analyzed to find the percentage of genes that possess H3K27me^3. For example, 252 genes are up regulated in *swn-7* the top 5% up regulated genes corresponds to 14 genes, 5 of which possess H3K27me^3.

Genotype	% up regulated	Term	Observed Frequency	Expected Frequency	p-value
swn-7	Top 5%	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
	Top 10%	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
	10-20%	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
	20-50%	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
	50%>	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
clf-28	Top 5%	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
	Top 10%	flower development	3 out of 19 genes	98 out of 27006 genes	0.00068
		post-embryonic development	3 out of 19 genes	180 out of 27006 genes	0.00417
	10-20%	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
	20-50%	iron ion transport	2 out of 22 genes	9 out of 27006 genes	0.00045
		zinc ion transport	2 out of 22 genes	12 out of 27006 genes	0.00083
		morphogenesis	4 out of 22 genes	299 out of 27006 genes	0.00184
		root morphogenesis	2 out of 22 genes	33 out of 27006 genes	0.00658
	50%>	There were no functional terms exceeding the p-value cut-off of 0.01.	N/A	N/A	N/A
swn-7 clf-28	Top 5%	embryonic development	15 out of 75 genes	108 out of 27006 genes	2.47E-20
		embryonic development (*sensu Magnoliophyta*)	13 out of 75 genes	89 out of 27006 genes	7.90E-18
		development	21 out of 75 genes	849 out of 27006 genes	3.15E-13
		lipid transport	6 out of 75 genes	99 out of 27006 genes	1.16E-05
		lipid storage	3 out of 75 genes	17 out of 27006 genes	0.00046
		vacuolar protein processing/maturation	2 out of 75 genes	4 out of 27006 genes	0.00154
	Top 10%	embryonic development	20 out of 156 genes	108 out of 27006 genes	6.68E-23
		embryonic development (*sensu Magnoliophyta*)	18 out of 156 genes	89 out of 27006 genes	2.87E-21
		development	36 out of 156 genes	849 out of 27006 genes	2.36E-19
		organogenesis	7 out of 156 genes	122 out of 27006 genes	0.00051
		lipid transport	6 out of 156 genes	99 out of 27006 genes	0.00174
		morphogenesis	9 out of 156 genes	299 out of 27006 genes	0.00459

swn-7 clf-28		lipid storage	3 out of 156 genes	17 out of 27006 genes	0.00872
	10-20%	There were no functional terms exceeding the p-value cutoff of 0.01.	N/A	N/A	N/A
	20-50%	regulation of transcription	46 out of 180 genes	1684 out of 27006 genes	6.38E-15
		transcription	46 out of 180 genes	1775 out of 27006 genes	4.64E-14
		nucleobase, nucleoside, nucleotide and nucleic acid metabolism	47 out of 180 genes	2623 out of 27006 genes	1.23E-08
		regulation of transcription, DNA-dependent	25 out of 180 genes	984 out of 27006 genes	6.37E-07
		transcription, DNA-dependent	25 out of 180 genes	1010 out of 27006 genes	1.06E-06
		physiological process	116 out of 180 genes	12552 out of 27006 genes	5.40E-05
		metabolism	92 out of 180 genes	9708 out of 27006 genes	0.00134
	50%>	regulation of transcription	27 out of 128 genes	1684 out of 27006 genes	9.36E-07
		transcription	27 out of 128 genes	1775 out of 27006 genes	2.77E-06
		regulation of transcription, DNA-dependent	17 out of 128 genes	984 out of 27006 genes	0.0002
		transcription, DNA-dependent	17 out of 128 genes	1010 out of 27006 genes	0.00029
		nucleobase, nucleoside, nucleotide and nucleic acid metabolism	29 out of 128 genes	2623 out of 27006 genes	0.00057

Table. 4. 3. Partitioning of up regulated genes and possessing H3K27me^3 of how highly up regulated, and analysis of over represented biological processes.

Genes found up regulated in *swn-, clf-28,* and *swn-7 clf-28* were partitioned into five sections based on how highly they were expressed compared to the total up regulated genes. These genes were then analyzed to discover over-represented biological functions. Over-represented biological processes were judged by GO term functions that found more commonly in a given gene list compared to the expected frequency in the genome, the cut-off of P = 0.01.

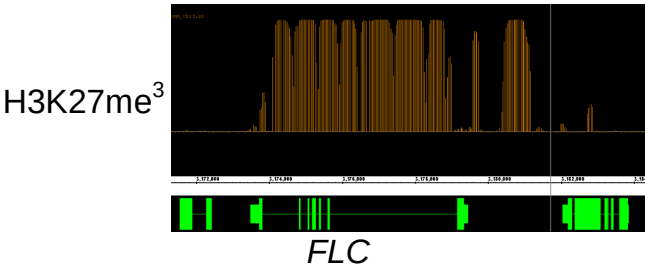

Fig. 4. 8. Enrichment of tri-methylation on lysine 27 histone 3 (H3K27me^3) at the *FLC* locus.
The Orange bars indicates the presence of H3K27me^3 with a 20 base pair resolution locus. Block vertical green bars represent exons of coding genes, and green horizontal lines represent introns. N.B. the H3K27me^3 marks are restricted to the *FLC* locus and not the contiguous genes. (Figure taken from (Zhang *et al.*, 2007).

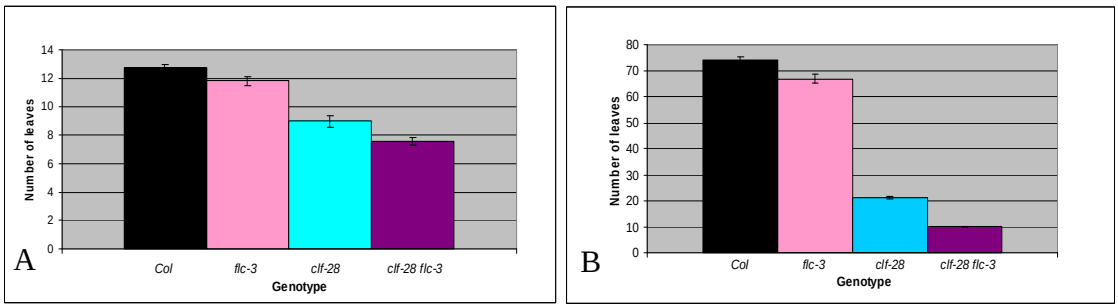

Fig. 4. 9. The flowering time in long and short days in wild-type, *flc-3, clf-28,* and *clf-28 flc-3.*
(A) Flowering time in long days scored by number of leaves. (B) Flowering time in short days. In both long and short days *clf-28 flc-3* plants flower earlier than either parent mutants, this effect is enhanced in short day conditions.

Fig. 4. 10. *clf-28 flc-3* mutants show an enhanced *clf-28* phenotype.
clf-28 flc-3 plants compared to *clf-28 FLC+* plants in short day conditions, showing earlier flowering and more severe leaf curling. (Scale =7 cm)

5. 0. Discussion: The dynamic Pc-G, processes it regulates, direct targets, and an insight into Pc-G mediated gene repression.

5. 1. Introduction

I set out my investigations with several objectives including, firstly, to identify a discrete role for *SWN* in development, secondly, to dissect the *swn- clf-* phenotype and uncover novel roles in development, thirdly, to resolve the discrete and redundant roles of *SWN* and *CLF*, and fourthly, to identify novel targets of the Pc-G. In this Chapter I will evaluate and discuss the evidence presented above involving transcriptional profiling and phenotypic characterization, with regards to the outlined aims. I shall also explore the limitations of these results, and some of the questions raised and the experiments required in the future to validate the derived hypotheses.

To uncover the role of the Pc-G, specifically *SWN* and *CLF*, in *Arabidopsis* development detailed phenotypic analysis of the *swn-* and *swn- clf-* mutants was performed. This revealed potential novel roles in development that are Pc-G regulated. To further expose the role of the Pc-G in development, identify potential targets, and explore the proposed redundancy microarray analysis was undertaken using *swn-*, *clf-*, and *swn- clf-* mutants. This gave insight into which genes are mis-expressed in the Pc-G mutants and consequently cause the observed phenotypes, indicating their roles in wild-type plants. The genes found mis-regulated in the mutants provided a platform to derive genome wide lists of potential direct targets of the Pc-G. This allowed the discovery of universal Pc-G regulated processes, and the dynamic nature of the Pc-G complexes, and in turn permitted insight into the potential mechanism by which the Pc-G mediates these processes.

5. 2. 0. *SWN* plays a discrete role in the juvenile to adult transition

The extensive phenotypic characterization of *swn-* mutants (Chapter. 1.) revealed a tendency to be delayed in the J-A transition, and to produce leaves faster than wild-type

plants in long and short days. These results were replicated multiple times in the *swn-3* allele, and were also observed in independent *swn-* alleles. The J-A transition appears to be a variable process, and some individual *swn-* plants analyzed showed a wild-type J-A transition and leaf initiation. Despite this, the population average differed from wild-type plant in each of the *swn-* alleles tested.

Possible genes that might result in either, or both, *swn-* phenotypes may be found in the type of genes found mis-regulated in *swn-7*. Many genes are found to be involved in metabolic processes, and hormonal responses. The disruption of these processes could conceivably cause the *swn-* defects, as they are likely to play a role in growth and differentiation. The best-characterized pathways of regulating leaf initiation rate and the timing of the J-A transitions are discussed below.

5. 2. 1. *SWN* is unlikely to act via the *miRNA* pathways to regulate the juvenile to adult transition

The potential role of *SWN* in the J-A transition is novel, but how does this integrate with previous knowledge of the vegetative phase change? The micro-RNA *miR156* which repress *SPL's*, and *TAS3* (a trans-acting *siRNA*) which acts to represses *ARF3* and *ARF4*, appear to be the central pathways regulating the J-A transition (Fahlgren *et al.*, 2006; Schwarz *et al.*, 2008; Hunter *et al.*, 2006). For example, *miR156* is highly expressed after germination and represses *SPL1-9*, *miR156* expression gradually decreases over time, and the subsequent de-repression of *SPLs* causes the change of juvenile to adult leaf forms (Schwarz *et al.*, 2008; Wu & Poethig, 2006; Wang *et al.*, 2008). *TAS3* is known to repress *ARF3* and *ARF4* in a similar manner except that *TAS3* expression is constant throughout development. The de-repression of *ARF3* and *ARF4*, in an unknown mechanism, causes the transition of from the juvenile to the adult phase (Hunter *et al.*, 2006; Fahlgren *et al.*, 2006). One possible explanation for the *swn-* phenotype is that *SWN* functions to repress the expression of the *TAS3* and/or *miR156*. Repression of these *miRNAs* results in the elevated levels of *ARF3* and *ARF4*, and *SPL1-9* transcription factors that promote adult leaf identity (Fahlgren *et al.*, 2006; Schwarz *et al.*, 2008;

Hunter *et al.*, 2006). If *SWN* were involved in these pathways, in the absence of *SWN* (*swn*-) elevated expression of these *miRNAs* would continue to repress *ARF3* and *ARF4* and *SPL1-9*, resulting in delayed J-A transition and increases leaf initiation rate (Fig. 5. 1). In support of this idea, it was found that *TAS3* and *miR156* both possess the H3K27me^3 mark making them possible targets for Pc-G action (Zhang *et al.*, 2007). Although the Affymetrix array used for the transcription profiling of *swn-7* did not include *miRNA* probes, the genes regulated by *TAS3* or *miR156*, e.g. *SPLs*, *ARF3* and *ARF4* were analyzed. However, no known components of the *miRNA* regulated J-A pathway were found mis-expressed (by at least two-fold) compared to wild-type seedlings. This may indicate that *SWN* does not regulate *miR156* expression in seedlings. However, this may be due to the age of the samples taken, as at the 10 day old seedling stage expression of *miR156* is at its peak and *SPL9* and *15* are scarcely transcribed at this point (Schwarz *et al.*, 2008). Perhaps *SWN* acts to repress *miR156* at a later stage in development. Elevated *miR156* expression in later development would repress *SPL3* expression causing a delayed J-A transition. This effect could be caused by the disruption of the threshold of the *miR156*: *SPL3* levels required for transition.

The quantification of *TAS3* and *miR156* in wild-type and *swn*- lines at a range of developmental stages would further test whether the *swn*- phenotype is a result of these *miRNAs*. To confirm if *SWN* directly and discretely represses *miRNA* expression ChIP analysis could be used to confirm A) whether *SWN* is bound to *miRNA* loci. B) The presence of H3K27me^3 mark at the *miRNA* loci, and C) if the H3K27me^3 mark is present at the *miRNA* loci is lost in the *swn*- mutants. Conversely, the *SWN*-Pc-G may not affect the J-A transition though *miRNA* pathway, which would indicate a novel mechanism for regulating the J-A transition.

5. 2. 2. Candidate genes that could mediate effects of *SWN* on the juvenile to adult transition

It has been shown that an altered metabolism for example, secondary metabolites and hormones (Miyoshi *et al.*, 2004) can cause increased leaf initiation rate, and J-A transition defects. In rice, *PLASTOCHRON1* encodes a cytochrome (*CYP*) known to be involved in metabolic processing; *pla1*- mutants show an increased leaf initiation rate. This is seen in one of its homologues in *Arabidopsis CYP78A5*. *cyp78A5*- mutants show J-A and leaf initiation defects, but the another *PLA1* homologue, *cyp78A7*- plants shows a no defects although *CYP78A5* and *CYP78A7* are known to act redundantly (Wang *et al.*, 2008). *CYP78A5* is found up regulated in *amp1*- mutants. *AMP1* encodes a glutamate carboxylase, which show an increased leaf initiation rate and premature J-A transition (Helliwell *et al.*, 2001). This suggests that *CYP78A5* and *AMP1* carry out metabolite processing which acts to repress leaf initiation rate and J-A transition in wild-type plants.

CYP78A7 is up regulated in *swn-7*, and therefore it is possible that this is a cause of the *swn*- phenotypes (Fig. 5. 1). *CYP78A7* has not been directly shown to be involved in phase transitions. Therefore, the *CYP78A7* may be involved in the same pathway as *SWN* to aid in regulating leaf initiation and J-A transition. However, *CYP78A7* does not possess H3K27me^3 in 10-day-old seedling, suggesting it is not a direct *SWN* target and its up regulation is a secondary effect.

5. 2. 3. Conclusions

The phenotypic characterization of *swn*- alleles revealed that *SWN*-Pc-G promotes the J-A transition and represses the leaf initiation rate. However, the *swn*- is subtle and there is a possibility that the observed effect is due to secondary mutation in the *swn*-3 background. In order to prove, unequivocally, the role of *SWN* in the J-A transition a complementation experiment should be carried-out. The *swn*- plants carrying a

transgenic wild-type copy of *SWN* should all show wild-type J-A transition and leaf initiation rate if *SWN* was entirely responsible for the phenotypic defects.

Transcriptional profiling of *swn-7* mutants has identified a candidate that could cause the *swn-* phenotypes. It is likely that the de-repression of direct targets causes the mis-regulation of metabolic activities, such as glutamate or hormone signalling. This in turn may cause the mis-regulation of genes required for regulating metabolites like the cytochromes (Fig. 5. 1). The increased metabolites may delay J-A transition and promote leaf initiation rate, as has been previously indirectly suggested (Miyoshi *et al.*, 2004; Ahn *et al.*, 2002; Wang *et al.*, 2008).

In conclusion, *SWN* had been discovered to play a role in regulating developmental timing although the exact mechanism is unclear presently. I have also demonstrated that *SWN* and *CLF* have probably diverged to function discretely. This is illustrated by the phenotypic data, is also illustrated in a number of genes found discretely mis-regulated in *swn-7* microarray data compared to *clf-28*.

5. 3. 0. *SWN* is not required for the vernalization response

There is strong evidence that the Pc-G acts to mediate the vernalization response, as mutants in the Pc-G member *VRN2* are insensitive to vernalization treatment in a vernalization requiring background (Gendall *et al.*, 2001), and the H3K27me^3 mark increases at the *FLC* locus to in response to cold treatment to induce its repression (Bastow *et al.*, 2004; Sheldon *et al.*, 1999; Gendall *et al.*, 2001). Therefore the *E(z)* class of genes are probably required catalyze *FLC* repression, particularly as *SWN* and *CLF* have been shown to interact with *VRN2* (Chanvivattana *et al.*, 2004). Previous data indicated that *CLF* acting discretely was not responsible for repressing *FLC* in response to vernalization, as *clf-* plants (or RNAi knock down of *CLF*) have a normal vernalization response in a FRI+ background (Wood *et al.*, 2006; Chanvivattana *et al.*, 2004). As *MEA* expression is confined to seed set, it was therefore suggested that *SWN* could have a discrete function to regulate *FLC* post-vernalization. Following

vernalization *swn-3/-3* plants in a *FRI+* background flowered at the same time as *SWN+ FRI+* plants; indicating that *SWN* does not have a discrete function in the vernalization response. *VRN2* was recently found in the same complex as *SWN* in a tap tag immuno precipitation (Crevillen, P. & Dean, C. 2008), but CLF was not found. This suggests that SWN is the preferred Pc-G member to mediate the H3K27me^3 mark in the vernalization response but that in *swn-* plants CLF can functionally replace SWN. This indicates that SWN and CLF can act redundantly to co-ordinate the vernalization response.

Due to the phenotypic deformity of the strong *swn- clf-* mutants, which do not flower or make leaves, it is technically difficult to confirm and quantify the redundant role of *SWN* and *CLF* in vernalization. It could, however, be strongly inferred by analyzing *FLC* expression and H3K27me^3 at the *FLC* locus in *swn- clf-* mutants in a vernalization requiring background. The increase in *FLC* expression and reduction in H3K27me^3 at the *FLC* locus would strongly implicate *SWN CLF* in regulation of the vernalization response. This could be verified more directly by analyzing the phenotypically weaker *swn-1 clf-50* double mutant, or using an inducible knock down of both *SWN* and *CLF* and analysing flowering time, *FLC* expression, and H3K27me^3 status at the *FLC* locus.

5. 3. 1. *SWN* promotes flowering in *FRI+* backgrounds

It was found that *swn-3 FRI+* flowered later than *SWN+ FRI+* when grown without vernalization treatment. The effect was observable, but some *swn-* plants did flower at the same time as some wild-type plants. This suggests that *SWN* may play a role in inducing flowering through the autonomous pathway in *FRI+* plants that have not been vernalized. There are four obvious explanations for how *SWN* could affect flowering through the autonomous pathway with dependency on *FRI* that are discussed below:

1) The repression of *FLC* is a very likely possibility, as *FRI* act directly to promote *FLC* expression, and flowering time is extremely sensitive to *FLC* expression with higher *FLC* expression plants leading to later flowering (Johanson *et al.*, 2000). Therefore, in *swn- FRI+* plants increased *FLC* would cause delayed flowering compared to *SWN+*

FRI+. However, the ecotypes used in this study are *fri-* and lack a strong positive regulation *FLC* expression, causing low expression *FLC,* and this may explain why no flowering defects are observed in *swn- fri-* lines as *FLC* expression remains low. Therefore, *SWN* could play a role in directly repressing *FLC* independently of vernalization (and *CLF*). Thus, the *FRI+ swn-* plants would flower later than *FRI+* plants because *FRI+* promotes *FLC* expression, and the mutated *SWN* could not repress *FLC*. The resulting high *FLC* expression would result in later flowering than observed in *SWN+*.

2) *SWN* could directly repress *FRI*. Thus, *swn-* plants would have elevated *FRI* expression; resulting in *FLC* up-regulation. The increased *FLC* expression would result in later flowering in non-vernalized plants. This effect is not seen in post-vernalized plants as *FLC* is down regulated by *SWN-CLF*-Pc-G. No flowering time defect occurs in *swn- fri-* plants as there is not positive regulation of *FLC*.

3) *SWN* could control expression of both *FLC*, and positive regulators of *FLC* such as *FRI*. This could also increase the regulatory control of *SWN* flowering time.

4) *SWN* may act in an *FLC* independent pathway but dependent on *FRI+* to regulate flowering time in plants that require vernalization.

Each of these explanations could also explain why *SWN* has been conserved. If *SWN* is acting discretely in natural accessions regulating *FRI* and/or *FLC* expression in the autonomous pathway it could contribute to the fitness of the plant by regulating flowering time. This hypothesis for *SWN* conservation cannot explain the conservation of a *SWN*-like gene in species that lack *FLC/FRI* orthologues, like rice (Tadege *et al.*, 2003).

To evaluate whether *SWN* had been conserved to repress *FRI* or *FLC*, several experiments are required. Firstly, there is a need to validate that *SWN* does repress *FRI* and/or *FLC*. This could be achieved by looking at *FRI* and *FLC* expression in *swn-*

plants in *FRI+* and *fri-* background. Then it would be possible to confirm that *FRI* or *FLC* are direct targets of *SWN* by analysing if the *FRI* and *FLC* loci possess the H3K27me^3 mark and if this is lost in the *swn-* plants, and whether this shows any correlation with flowering time. This would have to be carried out in a variety of accessions could suggest why *SWN* has been conserved.

SWN conservation in groups which do not possess orthologues of *FRI* or *FLC* may indicate that *SWN* has acquired different functions independent of flowering time regulation, or that *SWN* has retained the same function in repressing floral pathways or floral repressors but acting on unrelated genes. This could be tested by analyzing the enrichment of SWN binding at potential target loci in different species.

5 . 4. 0. Microarray suggests discrete targets of *SWN* and *CLF*

The quality of the array data was validated by the fact that previously identified targets were well represented. For example, *AG* was mis-regulated in *clf-28* microarray data, and *STM* was found mis-regulated in *swn-7 clf-28* double mutants and a range of other genes were also validated based on the array dataset such as *similar to PRE1* in *swn-7* and *swn-2* alleles. The microarray analysis of *swn-7, clf-28,* and *swn-7 clf-28* revealed the first direct evidence that *SWN* and *CLF* play unique, common, and redundant roles in regulating gene expression of direct targets. The potential direct targets represent a range of genes involved in diverse biological functions, and revealed that *CLF* and *SWN* play discrete functions in regulating gene expression which were significantly different to biological functions found regulated by *SWN* and *CLF* acting redundantly. This illustrates that the duplication event of the *SWN* and *CLF* resulted in divergence to both regulate distinct transcriptional profiles, where *CLF* regulates flowering time and floral organ identity, *SWN* regulates cell redox homeostasis, and *SWN* and *CLF* together repress embryonic traits, promote flowering, and regulate stem cell maintenance. The loci where *SWN* and *CLF* regulate common targets do not appear to define any particular processes or transcription profile. The transcriptional profiling also correlates with the difference in *swn-7, clf-28,* and *swn-7 clf-28* mutant phenotypes, as *clf-* plants and *swn-*

plants show different phenotypes, indicating they have developed discrete roles in development. The discrete role of *CLF* in development has been confirmed, because *SWN* is not interchangeable with *CLF*. Over-expression of *SWN* (*35s::SWN*) in *clf-* plants cannot rescue the *clf-* phenotype (Chanvivattana *et al.*, 2004). The discrete role of *SWN* is yet to be confirmed, which could be tested by the over expressing *CLF* in a *swn-* background. If *SWN* plays a discrete role in development it would be expected that the *swn-* phenotypes would still be observed.

5. 4. 1. Conclusions: *SWN* and *CLF* are not functionally interchangeable and have discrete roles

When two genes act redundantly, there is a tendency for one of the genes to lose its function, for example, *CAL* and *AP1* carry out redundant roles the inflorescence meristem and floral identity in the *Col* accession, but *CAL* has become mutated and has lost its function in the Ws accession (Alvarez-Buylla *et al.*, 2006; Kempin *et al.*, 1995). Alternatively, redundant genes can acquire discrete roles via changes in expression or proteins function, for example, in *Antirrhinum PLENA* (*PLE*) and *FARINELLI* (*FAR*) share a high degree of sequence similarity, but they have differences in their expression patterns (Davies *et al.*, 1999). When *PLE* and *FAR* are mutated they exhibit very different floral defects, however the *ple- far-* double mutants indicates that they act partially redundantly (Davies *et al.*, 1999). This suggests they have acquired discrete functions but have retained redundant functions in floral development. Their discrete roles in development indicate why both *PLE* and *FAR* have been conserved. This appears similar to the roles of *MEA, SWN* and *CLF* in *Arabidopsis. SWN, CLF* and *MEA* have striking sequence similarity (Chanvivattana *et al.*, 2004). *MEA* expression has been confined to the endosperm (Wang *et al.*, 2006), whereas *SWN* and *CLF* have virtually identical expression patterns and act throughout development (Goodrich *et al.*, 1997; Chanvivattana *et al.*, 2004; Schubert *et al.*, 2006). Therefore, the discrete roles of *SWN* and *CLF* in development must lie in differences in their protein structure, function, and perhaps changes in interaction partners. This could be confirmed by expressing chimera proteins, by creating proteins that are composed different domains of the SWN and CLF

proteins in the SWN or CLF proteins in the *swn- clf-*. For example, by expressing a CLF protein containing the SET domain of SWN instead of the CLF SET domain in *swn- clf-* mutants, then one would hypothesise three phenotypes could result. 1) if the SET domain was responsible of the difference in discrete roles then the *swn- clf-* mutants would display a *clf-* phenotype as the loss of *SWN* would be complemented by the SWN SET domain. 2) If the SET domain was not the domain responsible of the divergence of role then a *swn-* phenotype would occur as transgenic *CLF* chimera protein would complement the loss of *CLF* in *swn- clf-* mutants. 3) An intermediate phenotype may result; this would imply that SET domains are the partial cause of the divergent roles of *SWN* and *CLF*.

5. 5. 0. *SWN* and *CLF* are master regulators of plants development

5. 5. 1. Microarray analysis suggested at least 1000 direct targets of *SWN* and *CLF*

In flies, the Pc-G complex is known to regulate many targets, more than 100 targets have been identified based on ChIP on chip and polytene bands analysis (Tolhuis *et al.*, 2006; Zink *et al.*, 1991). Pc-G targets implied roles in patterning, stem cell maintenance and cancer suppression (Kanno *et al.*, 2008). This indicated that the Pc-G was of major importance in animal development, to maintain given cell fates. In plants, prior to these investigations, very few Pc-G targets were known. The *clf-* phenotype was largely a result ectopic expression of a single gene, *AG* (Goodrich *et al.*, 1997). This suggested that the Pc-G might not have a significant role in plant development. This is consistent with the observations that specific plant tissues have a less rigid fate determination, as certain plant tissues readily de-differentiated. This is based on tissue culture studies in which tissue types can be converted and grown into entire plants (Ikeda-Iwai *et al.*, 2002), and laser ablation of root stem cells induces differentiated neighbouring cells to de-differentiate to become stem cells (van den Berg *et al.*, 1995). The severity of the *swn- clf-* double mutant phenotype suggested the Pc-G might have targets that are more numerous and play a wider role in development than previously thought, but it was unclear what these targets were, or indeed how many. The combined use of

transcriptomics and published ChIP on chip data suggests that in *Arabidopsis SWN* and *CLF* have over one thousand redundant direct targets, the majority of which are novel, and many more secondary targets. However, the functional significance of most Pc-G targets has yet to be determined, but here I shall discuss the ones for which some genetic and phenotypic data that supports the role the Pc-G as a master regulator of many aspects of development.

5. 5. 2. *SWN* and *CLF* repress embryonic traits after germination

The similarity of the "pickle" root phenotype of *swn- clf-* and *pkl-* is striking, and *PKL* is a known repressor of embryonic traits, which acts partially though GA signalling to repress embryonic specific genes post-germination, potentially through chromatin remodelling (Ogas *et al.*, 1999; Li *et al.*, 2005; Henderson *et al.*, 2004). However, it was unknown if the similarity of *pkl-* and *swn- clf-* phenotypes was only superficial or whether it reflected a common cause. The transcription profiles of *pkl-* and *swn-7 clf-28* mutants revealed a significant proportion of genes commonly up regulated in both mutants and decorated with the Pc-G specific histone methylation. This overlap largely consisted of embryogenic specific genes. This demonstrates that *PKL*, *SWN,* and *CLF* act in a common pathway. I have shown that just like *pkl-*, the *swn- clf-* mutants also display accumulation of embryonic oils following germination.

When *hsl1- hsi2-* double mutants are grown on high sugar concentration they exhibit *swn- clf-* like phenotypes and display embryonic traits (Tsukagoshi *et al.*, 2007). Wild-type seedlings growth on high sugar concentrations also accumulate seed specific genes, such as *ABI4* expression (Arroyo *et al.*, 2003) but do not display a "pickle" root phenotype. When the weak *swn-2 clf-81* mutants were grown on high sugar concentrations they showed an enhancement of the root phenotype and increased phenotypic penetrance. This indicates that ectopic expression of genes specific to the embryo are a likely cause of the *swn- clf-* root phenotype.

In addition, many up regulated genes encode storage proteins and oils, and all of the master regulators of embryonic traits are up regulated in *swn-7 clf-28*. Most of which are direct targets of *SWN* and *CLF*, and many of these are also commonly up regulated in *swn- clf-* and *pkl-*, or *hsl1- hsi2-* mutants. This strongly indicates that the accumulation of embryonic specific traits is a partial cause of the *swn- clf-* root phenotype. This suggests that the fundamental role of the Pc-G is to repress embryogenic traits post germination (which had been suggested previously (Makarevich *et al.*, 2006)), but I extend and provide direct evidence for the this theory, and suggest that it does this partially through a common pathway with *PKL*.

A recent *PKL* microarray study revealed that genes de-repressed in *pkl-* such as SSPs, targets show a decrease in H3K27me^3 (Zhang *et al.*, 2008). This further indicates that *PKL* and the Pc-G are acting not only to repress genes in the same pathway, but that they probably achieve this through the same mechanism. This may suggest that PKL is either part of the PRC2 complex with SWN and CLF, or that it acts down stream of the PRC2 as a functional equivalent of the PRC1 in plants, potentially through chromatin remodelling thus reducing the accessibility of the transcription machinery. This may illustrate how the Pc-G mediates the stable repression of the embryonic specific genes throughout development.

The Pc-G appears to repress embryonic traits by directly repressing master regulators such as *FUS3, LEC2, ABI3,* and *ABI4* which all possess the H3K27me^3 mark (found in (Zhang *et al.*, 2007), and are the most highly mis-expressed genes in *swn-7 clf-28*. This strongly indicates they are direct targets of the Pc-G. In addition, CLF was found enriched at *FUS3* (Makarevich *et al.*, 2006). Not only are these master regulators direct Pc-G targets but the down stream targets of the master regulators such as the seed storage protein *EM1*, are also direct Pc-G targets as they also possess the H3K27me^3 mark. However, the confirmation that they are true direct targets is required by validating that their mis-regulation is correlated to a decrease in H3K27me^3 in *swn- clf-* mutants. These results indicate that several stages of the expression of the embryonic

trait pathway may be Pc-G regulated. The regulation of several stages of this pathway would to the stringency of repression of embryonic specific traits post-germination.

In conclusion, the *swn- clf-* phenotype is probably strongly influenced by the accumulation of master regulators of embryonic traits such as *LEC1*, *LEC2*, *FUS3,* and *ABI4,* and their down stream targets. However, the ectopic expression of embryonic traits alone is not sufficient to cause all aspects of the *swn- clf-* phenotype. Embryonic specific genes can only be partially responsible for the defects observed as the over-expression them not phenocopy the *swn- clf-* mutants. To test the role of the master regulators in the *swn- clf-* phenotype the triple mutants of *swn- clf- lec1-, swn- clf- lec2-,* and *swn- clf- fus3-* should be generated and their phenotypes characterized, and due to partially overlapping function of the master regulators, the quadruple and quintuple mutants should be analyzed. If the triple mutants alleviate the arrested development of the *swn- clf-* phenotype then they would be the likely cause of the *swn- clf-* phenotype and this may permit investigations into the other aspects of development like flowering time and flower development.

5. 5. 3. The Pc-G maintains the shoot meristem and leaf identity

In animals, the Pc-G has been implicated in maintaining stem cells in an undifferentiated state. It achieves this by repressing genes that promote differentiation (Chamberlain *et al.*, 2008). Many of these genes are bivalently marked, possessing both active and repressive histone modifications posed for future repression or activation in a cell type specific fashion as they differentiate (Pietersen & van Lohuizen, 2008). Once differentiated, genes promoting other differentiation states, stem cell fate and promoting cell proliferation, are maintained in a repressed state (Cao *et al.*, 2002). The Pc-G has been implicated in suppressing tumour development, as tumours cells show de-repression of embryonic stem cells and hypo-methylation of histone tails including H3K27me^3 (Metsuyanim *et al.*, 2008; Yu *et al.*, 2007). This suggests the Pc-G maintains differentiated cell fate by suppressing tumour development, by repressing embryonic stem cell fate through histone modification i.e. by repressing genes involved in stem cell

fate from being activated in differentiated tissues. Thus, the animal Pc-G is both required to maintain undifferentiated stem cells, and to maintain the given differentiation states.

The issue of stem cell regulation, meristem maintenance, and organization is raised in *swn- clf-* double mutants as SEM showed that organ formation was still occurring but in an irregular manner. Leaf-like structures are formed at the SAM up to the 21-day stage, but show deformities such as severe leaf serrations and root-like out growths. Genes involved in SAM maintenance, like *AIL5, AIL7, STM, LBD16*, and *CUC2* are up regulated in *swn- clf-* plants. This indicates that SAM is disrupted as judged by ectopic gene expression and morphology. Given this, it is interesting that the leaf-like organs formed in *swn-7 clf-28* after germination and appear to be initiated in a wild-type phylotaxis, suggesting that the SAM is intact transiently. The differentiation state of wild-type leaves is controlled by both auxin and transcription factors defining tissue differentiation (Barkoulas *et al.*, 2007). In *swn- clf-* double mutants leaves fail to maintain this initiated tissue fate. These leaf-like tissues show the most prominent phenotypic deformities of *swn- clf-* mutants. It is possible that leaf-like structures lose their differentiation state through the ectopic expression of genes required for maintaining the meristem. Ectopic *STM* has been shown to be capable of inducing *de novo* growth in young leaves (Hay & Tsiantis, 2006), so this seems a very likely cause of the abnormal outgrowths and serration on *swn- clf-* leaf-like structures. The combination of organ-like structures and outgrowths are probably a forerunner of the callus-like material of *swn- clf-* and *vrn2-1 emf2-3* mutants found later in development. The true meristem activity is lost in *swn- clf-* mutants over time i.e. *swn- clf-* plants continue to proliferate but fail to form organs, and this also occurs throughout the *swn- clf-* tissues to generate the callus-like material. This suggests that the Pc-G is required to restrict SAM function post-germination probably because of the increased/ectopic expression of the genes responsible for SAM identity (Fig. 5. 2).

This indicates the Pc-G plays significant roles in maintaining both SAM and leaf formation in wild-type plants, through repressing ectopic gene expression of genes that define the SAM, like *STM, AIL5,* and *AIL7,* outside of their expression domains (SAM).

This could be investigated further by analysis of SAM markers in the *swn- clf-* mutants and checking the H3K27me^3 state of these genes. Also, looking at the phenotype of the triple *swn- clf- stm-* mutant or mutant combinations with other genes known to regulate the SAM would reveal the role of SAM regulators on the *swn- clf-* phenotype. Over-expression of SAM regulators, such as *STM* and *WUS,* causes the formation of ectopic meristems and degenerative floral organs through the inhibition of cell differentiation (Lenhard *et al.,* 2002; Brand *et al.,* 2002; Gallois *et al.,* 2002). In addition, it also caused the formation of very large guard cells (Brand *et al.,* 2002). These phenotypic deformities are similar but not fully phenocopy *swn- clf-*mutants. This suggests that the *swn- clf-* mutant phenotype of the aerial parts is not fully attributable to the ectopic expression of *STM* or *WUS* specifically and so other genes are likely to influence the defects observed.

The other aspect of the meristem maintenance is meristem boundary regulation. A multitude of genes that regulate SAM size and spacing of lateral organs that are up regulated in *swn-7 clf-28,* including the *CUC2*, *LBD16,* and *LBD40* genes, indicating that both the SAM and lateral organ boundaries are directly regulated by the Pc-G.

What could this mean in the "bigger picture"? Meristem maintenance is essential for proper development in *Arabidopsis*, as plants lacking shoot meristem are not viable e.g. null *stm-* plants (Long *et al.*, 1996). *STM* is a direct target of the Pc-G and so regulates its expression in *Arabidopsis*. *STM* expression at leaf margins appears to be essential for species with compound leaves, as *STM* promotes the out-growths that develop into leaflets (Hay & Tsiantis, 2006). Compound leaf form has been adopted by many divergent species to gain selective advantage for the environment they inhabit e.g. palm species have longer petioles possibly to allow access to air and light in tropical climates. Changes in how and /or where the Pc-G regulates *STM* expression, may have allowed changes in the leaf form perhaps contributing to the evolution of compound leaves. In addition, the functional equivalent of *STM* expression is found in leaf margins and is required for somatic embryo production in asexually reproducing species such as *Kalanchoe* (mother of thousands) (Garces *et al.,* 2007). Therefore, small changes in the

way the Pc-G regulates *STM*, and possibly other genes, in species such as *Kalanchoe* may provide an evolutionary basis of the mechanism to form somatic embryos for asexual reproduction, particularly as the homologue of *LEC1* is often required for the correct somatic embryogenesis, also a Pc-G target (Garces *et al.*, 2007).

The role of the plant Pc-G in repressing the stem cell identity in differentiated tissue is reminiscent of the animal Pc-G and perhaps analogous to tumour suppression in animals, whereby the Pc-G represses embryonic stem cell fate in differentiated tissues (Yu *et al.*, 2007; Kanno *et al.*, 2008; Pietersen & van Lohuizen, 2008; Ben Porath *et al.*, 2008). Indeed the unregulated cell proliferation and reduced H3K27me^3 status found in *swn-3 clf-50* mutants (Lindroth *et al.*, 2004) is reminiscent of tumour cells (Metsuyanim *et al.*, 2008). This suggests a potentially similar role of the Pc-G in both plants and animals to control cell proliferation in differentiated tissues.

5. 5. 4. The Pc-G maintains root meristem identity

Root tissue was of particular interest as the *swn- clf-* mutants show a radically stunted "pickle" root tip phenotype and root tips that exhibit cells that were smaller and irregular in shape and orientation compared to wild-type roots. These observations brought about the hypothesis that the Pc-G is required to maintain the RAM tissues. To test if *swn- clf-* roots had lost stem cell maintenance *swn- clf-* roots were stained with Lugol's stain to observe any changes in differentiation in the root tip. This revealed a loss of the stem cell population in the root tip at 4 days old. Interestingly, the root cell files can still be observed after 14 days in *swn- clf-* mutants, at least 10 days after the loss of the stem cell niche.

The QC in the root tip preserves and choreographs the stem cell population, and essentially defines root identity and growth potential. The stem cells then give rise to the cell files that acquire a differentiated state (Sablowski, 2007b). There are several known mechanisms required for maintenance of root identity, or proper QC regulation. For example, *SHR* and *SCR* transcription factors are specifically required for endodermis

and cortex differentiation and maintaining the stem cell population through regulation of the QC (Helariutta *et al.*, 2000; Nakajima *et al.*, 2001b). Auxin distribution is also well documented as being required to maintain the RAM identity (Grieneisen *et al.*, 2007).

There was decreased expression of *SHR::GFP* and *SCR::GFP* after 14 days growth in *swn- clf-* and *vrn2-1 emf2-3* mutant, but they were comparable to wild-type at 4 days old. Expression of *Dr5::GUS* and *Dr5*::*GFP* was found to be similar to wild-type at 4 days and was reduced or lost at 14 days in *swn- clf-* and *vrn2-1 emf2-3* root tips. The reduction in auxin gradient, and *SHR* and *SCR* in the root tip occurred after the loss of the stem cell population, suggesting that the reduction in auxin nor the *SHR-SCR* pathway are not the direct cause of the loss of RAM function and differentiation in *swn- clf-* and *vrn2-1 emf2-3* mutant roots.

As the reduction of *SHR-SCR*, or auxin pathways are probably not responsible for the lost QC and stem cell identity, the question remains: what are the causes of the lost RAM identity? The mis-expression data for *swn-7 clf-28* revealed that very few genes required for the maintenance of the root cell population were mis-regulated, perhaps due to sensitivity issues. A known gene required for root stem cell maintenance which was down regulated was *WOX5*. *WOX5* is expressed from early embryogenesis, and its expression is confined to the QC post-germination (Haecker *et al.*, 2004). It is unlikely that the *swn- clf-* phenotype is solely due to the down regulation of *WOX5*, as although *wox5-* mutants show reduced root growth, indicating a role in QC differentiation (Sarkar *et al.*, 2007), they do not show the "pickle" root phenotype of *swn- clf-* mutants. It would be none the less interesting to test whether the over expression of *WOX5* in the QC could be capable of restoring the root stem cell population in *swn- clf-* mutants. However, it seems more likely that the *swn- clf-* stem cell population may be losing their indeterminacy through mis-regulation of genes that maintain QC identity other than *WOX5*. It is possible that the Pc-G is responsible for the maintaining RAM by repressing genes in the stem cell region that promote differentiation. When Pc-G function is disrupted genes not normally expressed in the RAM are ectopically expressed, such as floral/leaf homeotic transcription factors, and embryonic specific traits, these genes may

be responsible for the root stem cells acquiring determinacy in *swn- clf-* mutants. This may suggest a network whereby the stem cell maintenance genes are subordinate to factors inducing differentiation. If this is true, it could be suggested that the root has lost its identity, and acquired an embryonic maturation expression profiles, and because of the phenotypic similarity of *swn-7 clf-28* to *pkl-* "pickle" roots, as the *pkl-* root is that is caused by the de-repression of embryonic traits (Ogas *et al.*, 1999).

This is reminiscent of the animal Pc-G that has evolved to maintaining stem cell identity from gaining a differentiation state in embryos by repressing genes that promote differentiation (Lee *et al.*, 2006). However, it is presently unknown whether the bivalently histone marks genes that promote cell differentiation during embryogenesis in *Drosophila* embryos, are found in the root meristem population or whether they are used in the same manner.

5. 5. 5. Conclusion: The Pc-G is required to maintain root and shoot meristems

It appears that both the RAM and the SAM are Pc-G regulated as their integraty is lost, judged by phenotype in the *swn- clf-*, probably due to the loss of regulation through the H3K27me^3 mark. The RAM and SAM are formed during embryogenesis (Jenik *et al.*, 2007) and *SWN* and *CLF* do not appear to be required for normal embryo development. The loss of meristem activity in *swn- clf-* seedlings does not occur immediately after germination, and requires time, and probably cell division. This suggests two points: 1) the phenotype of both the "pickle root" and the irregular leaf-like structures are a consequence of loss of differentiation in newly divided cells and tissues. 2) There is a transient maintenance of the RAM and SAM independent of *SWN* and *CLF* post germination. The QC does not divide as frequently as the stem cell population that it regulates (Dolan *et al.*, 1993). This may explain why the *de novo* tissues exhibit defects and why the RAM and SAM pluripotent state is maintained longer than the frequently actively dividing tissues**.** In essence, it appears that the SAM and RAM are maintained by the Pc-G, but the processes mis-regulated in *swn- clf-* in the root and shoots are likely to be different, as the phenotypic deformities suggest. This results in the different

phenotypes in early development in *swn- clf-* in the aerial and below ground parts. Judging from the phenotypic evidence and known roles of *swn-* and *clf-*, the "pickle" root phenotype is largely due to the de-repression of embryonic traits, and the aerial defects are probably caused by de-repression of genes required for SAM maintenance, and floral organ identity. Therefore, the Pc-G appears to maintain RAM tissue fate by repressing genes defining differentiated tissue fates from being ectopically expressed in the RAM tissues, and regulate SAM fate by repressing SAM specific genes from being expressed in differentiated tissue like leaves (Fig. 5. 2).

These opposing methods of maintaining tissue fate is reminiscent of the way in which the animal Pc-G has evolved to repress genes in stem cells that promote differentiation in embryos, and to maintain differentiated tissue by repressing embryonic stem cell fate gene expression. The difference in maintaining tissue fate occurs in different stages of development. Conversely, the *Arabidopsis* Pc-G both the maintain stem cell and differentiated tissues types simultaneously and this plainly exemplifies the major differences in growth strategy of plants and animals, but also illustrates that the Pc-G complex in its divergent forms have evolved to regulate similar processes in development.

5. 5. 6. A mechanism for Pc-G mediated repression

Removing Pc-G function de-represses embryonic traits and this reveals a new prong in the multi-pronged approach to repress embryonic traits post-germination. Mutations of Pc-G, *PKL*, *BRM,* or *HSI2 HSL1,* or inhibition of histone acetylation by pharmacological treatment (TSA) (Tanaka *et al.*, 2008), or response to sucrose (Arroyo *et al.*, 2003) all result in the de-repression of embryonic traits. These include both master regulators of embryonic traits like *FUS3, LEC1, LEC2, ABI3,* and *ABI4*, and their targets such as seed storage proteins (SSPs) and late embryogenesis abundant proteins (LEAs). Genes mis-expressed in these mutants and conditions revealed significant overlaps with genes found up regulated in *swn-7 clf-28* suggesting that all these processes link with the Pc-G to regulate embryonic repression post-germination.

Considering the probable interactions of genetic and epigenetic mechanisms that regulate the repression of embryonic traits in the post-germinative root, I will use this to discuss the possible *modus operandi* of Pc-G (Fig 5. 3). The *SWN/MEA-FIE-FIS2-MSI1*-Pc-G (FIS2 complex) acts in the endosperm to repress *FUS3* by creating/maintaining the H3K27me^3 mark. Both master embryonic regulators and embryonic traits are expressed in the embryonic root during embryogenesis. This activation is correlated with "active" histone marks like acetylation and probably H3K4me^3 marks. As the embryo develops into the dormancy phase late in seed development, embryonic maturation traits are repressed. The "active" histone acetylation marks are subsequently removed by histone deacetylases (*HDA6* and *HDA19* (Tanaka et al., 2008) from both master regulators and their down stream targets. The removal of the histone acetylation mark may act as a "signal" for histone methylation to occur, for example by removing an inhibitory mark (which has been previously observed in animal systems (Ikura *et al.*, 2007)). The *SWN/CLF-FIE-EMF2-VRN2- MSI1*(?)-Pc-G (EMF2 complex) creates and maintains the H3K27me^3 mediated repressed state of the embryo maturation genes, during the desiccation stage of embryogenesis through out the embryo. This repressed state is maintained throughout development, from embryo to seed set, by the EMF2 complex. The Pc-G mediated mark is likely to be interpreted by a functional equivalent of the PRC1 protein/complex, to stably maintain the H3K27me^3 mark, either by changing the chromatin nature (by creating localized regions of "closed" chromatin or pseudo-heterochromatin (for example *PKL*)), or by protecting the H3K27me^3 from active removal by histone demethylases, or passive removal. These two approaches excludes the transcriptional machinery from transcribing the silenced loci, to provide a rigorous method to repress gene repression, like embryonic traits through development.

The epigenetic regulation in addition to further transcriptional regulation inhibits the action of transcription factors and hormonal signalling from activating embryonic traits. *HSL2 HSI1* act redundantly to prevent sucrose (or sucrose signalling) from activating embryonic traits post-germination. *PKL* is known to repress embryonic specific gene

expression in the same pathway as GA signalling. However removing any component of this proposed mechanism results in ectopic embryonic traits, illustrating that either sucrose itself, or sucrose regulated transcription factors can stimulate expression of embryonic traits. This could occur by disrupting the PRC1-like/ PRC2-H3K27me^3 interaction to allow access to the transcriptional machinery. It is also possible, if unlikely, that sucrose or sucrose signalling could open "closed" chromatin resulting in transcriptionally active chromatin independent of H3K27me^3.

5. 6. The curious incident of the *swn- clf-* embryos

The *mea-* , *mea- swn-*, *fis2-*, *fie-*, and *msi1-* mutants without fertilization all show autonomous seed set or aberrant proliferation of embryo and endosperm traits (Luo *et al.*, 2000; Kohler *et al.*, 2003a; Wang *et al.*, 2006; Chaudhury *et al.*, 1997; Guitton *et al.*, 2004). This provides evidence that a MEA-FIS-Pc-G complex represses endosperm development prior to fertilization (Luo *et al.*, 2000). The finding that *swn-* enhanced the *mea-* phenotype suggests that *SWN* and *MEA* act with partial redundancy prior to fertilization (Wang *et al.*, 2006). *SWN* and *CLF* also act redundantly in post-germination development (Chanvivattana *et al.*, 2004). *SWN* and *CLF* are both expressed during embryogenesis, but neither *swn-* or *clf-* single mutants show embryo defects, which indicated that *SWN* and *CLF* might act redundantly at this stage. However this appears not to be so, as *swn- clf-* embryos are morphologically wild-type. This suggests that *SWN* and *CLF* do not act redundantly with each other during embryogenesis. *swn-2 clf-81* mutants show ectopic *MEA* expression post germination (Jullien *et al.*, 2006). This could imply that ectopic *MEA* expression in the *swn- clf-* embryos may mask the effect of losing *SWN* and *CLF* function, resulting in morphological normal embryos. However, there is no evidence that *MEA* can fulfil the roles of *SWN* or *CLF*. *MEA* does act partially redundantly with *SWN* during embryogenesis, and *SWN* acts redundantly with *CLF* post germination. It is therefore possible that *MEA* could carry out some the SWN/CLF-Pc-G functions if ectopically expressed post germination, even transiently, could result in Pc-G targets acquiring of H3K27me^3 resulting in relatively normal gene expression for a short period as is seen in *SCR::GFP* and *SHR::GFP*. Subsequent

dilution of histone marks, perhaps through cell division in the absence of *SWN* and *CLF* may potentially explain the time lag of the *swn- clf-* phenotype and the increasing severity over time. This could be tested by quantifying methylation status, using a florescent anti-H3K27me^3 antibody and comparing the daughter cell to the progenitor cell in a *swn- clf-* background to confirm that the H3K27me^3 is lost passively by dilution through successive rounds of cell division in the absence of histone methyltransferase activity. *WOX5* (described above) may be an ideal model for these experiments as it is expressed during embryogenesis and post-germination specifically in the QC.

If the ectopic *MEA* expression in the embryo does suppress the *swn- clf-* embryo phenotype, then an enhanced *swn- clf-* phenotype may be seen in *swn- clf- mea-* triple mutants. However, this is technically challenging due to the severity of the *swn- clf-* mutant phenotype and the maternal effect early embryo lethality of the *mea-* phenotype. Thus, it is impossible to observe the effect of the triple mutant later in development. This, however, could be analyzed by the inducible *siRNA* knock down of *MEA* in a *swn- clf-* background at different stages in seed and post germinative development. If *MEA* was responsible for suppressing the *swn- clf-* phenotype, then it would be predicted that further suppression would results from over expressing *MEA.*

It is curious though that the none of the *E(z)* homologous are required directly for maintaining the early patterning of the embryo, but they are required to maintain the given pattern after germination. This could be similar to what occurs in the *Drosophila* embryo, where embryonic segments are initially patterned and determined by the pair-spacing and gap genes regulating *HOX* gene expression patterns, which are only subsequently maintained by the Pc-G complex throughout development (Francis & Kingston, 2001). A similar mechanism maybe employed in plants, where the basic body plan is determined in embryogenesis initially by auxin distribution to form asymmetric cells but then is patterned and co-ordinated by transcription factors independent of the Pc-G. However, at present how these patterning transcription factor are regulated in embryogenesis is a mystery. However, it is possible that the Pc-G may not be required

until germination to maintain the prescribed expression patterns or repressed state of transcription factors.

MEA and *SWN/CLF* both repress embryonic traits, as shown by the *mea-* and *swn- clf-* mutants exhibiting a broad de-repressed seed maturation profile. This may indicate that this is the ancestral role for the Pc-G. It is believed, based on protein homology and the similarities of the surrounding chromosome regions, that *SWN* was the ancestral precursor of *MEA*, and *SWN* also gave rise to *CLF* (Spillane *et al*., 2007). This indicates that the ancestral *SWN* was used to repress embryonic specific traits. *MEA* expression may have become confined to seed development, and its homologues, *SWN* and *CLF* continued to repress embryonic traits post germination. *SWN* and *CLF* then gained further functions with respect to meristem maintenance, floral organ identity, metabolism, plastochron rate, and the J-A transition.

The duplication of *E(z)* proteins appears to have occurred in other Pc-G components, like *EMF2* and *VRN2*. This duplication of the Pc-G components seems to have allowed the expansion of roles, therefore when the redundant functional components are removed it results in an almighty disruption of many processes that effect development and growth. Analysis of mutants in Pc-G genes in species which are believed to be ancestral or pre-date *Arabidopsis* speciation such as mosses and monocots would be insightful in resolving this question. If “embryonic traits” were mis-expressed this would be strongly indicative that the ancestral function of the Pc-G in plants is to regulate embryonic development both during and post-embryogenesis.

5. 7. The Pc-G regulates the expression of *AGAMOUS-Like* genes

In *Drosophila* the PRC2 and PRC1 are required for the differentiation of segments in the *Drosophila* embryo by spatially regulating gene expression pattern of *HOX* gene clusters (Scott & O'Farrell, 1986). The *HOX* genes are found in clusters interspersed with regulatory sequences, including PREs. PREs are a sequence specific region found to be a binding site for the Pc-G complexes, and which are tri-methylated at lysine 27. PRC1

binds to the H3K27me^3 mark to stably maintain the repressed state. Mutating Pc-G members in *Drosophila* results in the ectopic expression of *HOX* genes (Dejardin & Cavalli, 2005). This is reminiscent of what is found in *Arabidopsis* . In *clf-* mutant there is a de-repression or ectopic expression, in space and time of *AG* (Goodrich *et al.*, 1997). A range of other *AGL* genes are found up regulated in both *clf-28* and *swn-7 clf-28*. This suggests that, like in *Drosophila*, the plant Pc-G complex regulates a developmentally important gene clade to regulate differentiation. The *AGL* genes are grouped into three mains clades of *AG*, *AP3*-like, and *AP1*-like (Favaro *et al.*, 2003). Several further genes that do not fall into these clades. Each of these clades of *AGL* genes are found up regulated in *swn- clf-* and nearly half of the *AGL* in the genome possess H3K27me^3, suggesting the ancestral *AGL* gene was Pc-G regulated. The *AGL*s do not appear to be present in clusters like the *Drosophila HOX* genes, and the MADs-box transcription factors (like the *AGL* clade) are unrelated to the *HOX* genes as they lack the homeobox motif that characterizes the homeotic genes. This indicates that the plant PRC2 and the fly PRC2 have been recruited to their respective targets independently during the evolution of plants and animals. The conservation of the PRC2 is evident with the PRC2 being present in both plants and flies, but the lack of PRC1 in plants also underlines the major differences in animal and plant Pc-G based gene repression.

5. 8. Default silencing of *FLC* by the Pc-G

The Pc-G complex mediates the vernalization response (Sheldon *et al.*, 2000). In plants requiring vernalization, i.e. possessing functional *FRI* genes, *FRI* activates high *FLC* expression that inhibits flowering until they undergoe vernalization (Johanson *et al.*, 2000). During vernalization the Pc-G is recruited to *FLC*, which gains H3K27me^3 causing its repression and permitting flowering to occur (Wood *et al.*, 2006; Gendall *et al.*, 2001; Bastow *et al.*, 2004). However, the *FLC* locus in *fri-* plants possesses H3K27me^3 (Zhang *et al.*, 2007; Turck *et al.*, 2007), suggesting that the Pc-G regulates *FLC* without the requirement for vernalization, in addition, the array data for *clf-28* and *swn-7 clf-28* (in a *fri-* background) show an up regulation of *FLC*, which was validated by Q-PCR analysis of *FLC* expression levels. As the samples tested were in a non-

vernalization requiring background it seems likely that the Pc-G does regulates *FLC* expression independently of vernalization. The significance of this was analyzed by observing flowering time of *clf- flc-3* (Col), which was earlier than *clf-28 FLC+* plants. This shows that although *clf-28 FLC+* plants flower early this is some extent mitigated by up regulation of *FLC*. This surprisingly indicates that in early flowering accessions (*fri-*) the Pc-G has a role in promoting flowering by repressing *FLC*. Conversely, the Pc-G plays a role in delaying flowering by repressing flowering activators such as *FT, SPL3,* and *AG* independently of vernalization. This demonstrates the duality of the Pc-G in flowering by regulated antagonistic flowering regulators. This raises the possibility that in *fri-* lines silencing of *FLC* by the Pc-G may be the default state. This would suggest that the role of *FRI* could be to inhibit the Pc-G from silencing *FLC*. This hypothesis could be tested by analysing the relative enrichment of *CLF* and *SWN* at the *FLC* locus in both *FRI+* and *fri-* plants. If FRI+ did in some manner inhibit *SWN* and *CLF* from accessing the *FLC* locus then less enrichment of the *SWN* and *CLF* at *FLC* would be expected in the *FRI+* lines.

5. 9. A Gradient of Pc-G function reflects its role in regulating development

The mis-expression found in *clf-28* and *swn-7 clf-28* microarray were compared to mis-expression founds in other Pc-G mutants, *emf1-*, *emf2-*, and *msi1-*. This was carried out as is would be expected that there would be mis-regulation of common targets as they act in the same complex. It is significant the that genes found up regulated in *clf-28* or *swn-7 clf-28* and *emf2-* or *msi1-* showed a vast enrichment of genes that possessed the H3K27me^3 mark, compared to the single mutants. Thus is consistent with *EMF1, EMF2* and *MSI1* being members of a Pc-G complex and act to repress common target gene expression through the H3K27me^3 mark. It is also striking that the mutant datasets do not completely overlap. Indeed there was relatively little overlap between mutants. This suggests that the relatively small changes in the level of Pc-G function can make major differences in which targets become mis-regulated. This is reflected in the phenotypes of the mutants. This is supported by further genetic evidence, relatively small differences in *EMF2* activity can cause differenced in flowering time, *emf2-3,* and *emf2-10* mutants are

flower earlier than *clf*- alleles (Fig. 5. 4). The *emf2-3* mutation is null, whereas, the *emf2-10* is a point mutation and *emf2-3* is phenotypically much stronger than *emf2-10* i.e. earlier flowering and more severe floral defects (Fig. 5. 4). The *vrn2-1 emf2-3* double mutant shows a gross disruption of developmental processes; however, this is phenotypically very different to the *vrn2-1 emf2-10* mutants, which are comparatively healthy but are late flowering (Fig. 5. 4). These variable and unpredictable phenotypes of Pc-G mutants is probably a cause of several factors firstly, the number and type of genes mis-expressed**,** the level and timing of ectopic expression in accordance with the level of Pc-G function and therefore the mutant phenotypes observed. Secondly, the Pc-G regulates targets of developmental pathways that act antagonistically (e.g. *FLC* and *FT*) implies that the resulting phenotype of Pc-G mutants is a consequence of which targets are over-expressed. For example, the late flowering of *vrn2-1 emf2-10* may be a consequence of extremely high expression of *FLC*, although other genes acting antagonistically in the same pathway may also be mis-regulated. Thirdly, Pc-G targets have different degrees of redundant regulation, for example, *FLC* is much more highly up regulated in *swn-7 clf-28* than in *clf-28,* but *AG* expression in *clf-28* is higher than *swn-7 clf-28*. This suggests that *FLC* is more redundantly regulated by *SWN* and *CLF* than *AG*. Hence, when varying Pc-G function is removed the mutant phenotype is at mercy of which genes are mis-expressed which may depend on unpredictable gene mis-expression and antagonism in developmental pathways (Fig. 5. 4).

5. 10. Concluding remarks

The role of the Pc-G in development is undeniable given the dramatic and grotesquely disorganized phenotypes of mutants with disrupted Pc-G function. This is likely the consequence of the Pc-G in maintaining differentiated tissues, the pluripotent nature of meristematic tissue, and by regulating developmental transitions, such as flowering and the J-A transition. The Pc-G proteins SWN and CLF have accrued discrete and redundant roles in plants development, through changes in their protein function. The phenotypic consequence of removing Pc-G activity appear to depend on the degree of Pc-G function removed, as the Pc-G mutants have variable phenotypes, this is probably

a consequence redundancy of some Pc-G components, which targets are mis-expressed, and the mis-regulation of targets which act antagonistically in the same developmental pathways. The genes found decorated with H3K27me^3 were found to be the most strongly over-expressed in *swn-7 clf-28* this is consistent with them being direct targets repressed by the Pc-G complex. There are over one thousand potential direct Pc-G targets, indicating that the Pc-G is a general regulator of plant transcription that acts to co-ordinate development. This is reminiscent of the animal Pc-G, as is the roles in maintaining stem cell and differentiation states. However, cell fate in animals is more rigidly maintained compared to plants, which are more flexible. This suggests that the effects of the plant Pc-G on development are readily reversible. This may indicate that a molecular basis for plant growth plasticity entwines the regulation of the Pc-G and the histone demethylases particularly, H3K27me^3 demethylases. I propose this on the basis that the Pc-G acting through histone methylation is required to co-ordinate cell fate, that the removal of H3K27me^3 marks maybe required to co-ordinate the reversal of cell fate.

5. 11. 0. Future work to discover the role of Pc-G in other aspects of development and the mechanism of Pc-G action

The analysis described and discussed above has thrown light onto the many and varied functions of the Pc-G, potential targets, and possible insight into the mechanistic basis of the Pc-G. However, many pertinent, and fundamental, questions have arisen subsequent to these investigations, some of the most pertinent questions are discussed below.

5. 11. 1. Validation of Pc-G targets and the discovery of unique, common, and redundant Pc-G member functions

There is a requirement to fully validate the scope of Pc-G in gene regulation and find to potential independent roles for the members of the Pc-G outside Pc-G function. This could be achieved by ChIP on chip and microarray analysis. Comparing the location of binding of FIE, MSI, SWN, CLF, VRN2, and EMF2 in the genome would provide a precise map of Pc-G targets. Comparing potential binding loci of the Pc-G with the wild-

type epigenome map of H3K27me^3, would strongly imply which loci were direct targets of the Pc-G. This could be validated by ChIP on chip analysis of the H3K27me^3 mark in Pc-G mutants such as *swn- clf-* and *fie-* represent the total Pc-G function. Those genes identified as being bound by multiple members of the Pc-G and possessing H3K27me^3 in wild-type plants, which lose this mark in *swn- clf-* and *fie-* plants would be direct targets. This would plainly show that the Pc-G was bound at those loci in order to make or at least maintain the H3K27me^3. The comparison of these potential direct targets to gene expression data would show the functional role of the H3K27me^3 mark. All of these experiments should be carried out at the same developmental stage and same tissue type, for example whole seedlings at 10 days old, as it is highly probably that the Pc-G has many targets at different times in development, and these probably vary in different cell types. This approach would also reveal the potential role of Pc-G components acting inter-specifically in the Pc-G i.e. the discrete functions of the components, and potentially independently of the Pc-G, and H3K27me^3, and thus the common and discrete functions and targets of the homologous members of the Pc-G.

The potential Pc-G target loci could be compared to other epigenome maps of histone marks like other methylation and acetylation marks; this may give insight in to the mechanistic function of the Pc-G and the involvement and interaction of other histone modifiers in development. If localizations are found then introducing a mutant histone acetlyases and histone deacetylases into the *swn- clf-* background to verify the effects on gene expression and phenotype.

5. 11. 2. Are histone demethylases essential for the flexible cell fate in plants?

I proposed that the Pc-G is required to maintain the cell fate of differentiated tissues post-germination; therefore, histone demethylases may be required for the plant cells to lose differentiation fate. This could be investigated by assaying the ability of plants with mutated histone demethylases to regenerate or to de-differentiate. A neat model to test this would be to laser ablate the root stem cell population and to then analyse the root cell division, as wild type plants de-differentiate cells surrounding the stem cell niche to

form a new stem cell population, and can continue to generate new cells (van den Berg *et al.*, 1995). Therefore, if histone demethylases play a role in de-differentiation then histone demethylases mutants would show reduced ability to carry out this process.

5. 11. 3. Does the Pc-G have novel functions in defence response?

Those genes up or down regulated with H3K27me^3 in *swn-7*, *clf-28*, and *swn-7 clf-28* mutants are considered as potential direct targets. They frequently represent genes involved in response to stress and pathogens. This suggests that the Pc-G plays a role in the defence response. It was previously unknown whether histone modification played any role in the defence response. It is known epigenetic variation of resistance gene loci and clusters, in the form of DNA methylation, affects gene expression and pathogenic resistance (Stokes & Richards, 2002; Yi & Richards, 2008; Stokes *et al.*, 2002). In theory, epigenetic regulation by histone modification is ideal for regulating the defence response, as defence response genes require repression until pathogenic attack. However, their expression state would have to be "reset" in the following generation or after pathogen attack; otherwise, it would confer a reduced state of fitness to the plant, as is seen in mutants that have constitutively active pathogen defence responses. For example, the constitutive expression of *SSI4* results in necrotic lesions and a severely dwarfed phenotype (Zhou *et al.*, 2008). This could be analogous to the mechanism that Pc-G mediates of the vernalization response, where the high *FLC* is repressed by the Pc-G to permit flowering, but the next generation vernalization is once again required to permit flowering, therefore suggesting that the Pc-G mediated repression is some how "reset" to allow high *FLC* expression.

To investigate this possibility, challenging *swn-*, *clf-*, and *swn- clf-* mutants with both pests and pathogens, could reveal if the Pc-G does contribute to the defence response. If the Pc-G does play a role in the defence response then the over-expression of genes involved in pest/pathogen defence could confer increased resistance the plant. This process may be similar to the regulation of vernalization response. If so then this may provide an efficient model to investigate the "re-setting" process, some defence response

genes are known to be expressed only for the duration of attack (Chassot *et al.*, 2008). These may be epigenetically re-set following attack, and not in the next generation, until induced by pests or pathogens.

4. 11. 4. Discovering the role of the ancestral Pc-G and the role of *SWN*

Analysis of the Pc-G members in evolutionarily distant species may also provide evidence for the ancestral Pc-G function. The conservation of a *SWN*-like gene in evolutionarily distant species does raise the question of what roles *SWN* may play in their development. Perhaps analysis of homologous *swn*- mutants in mosses, grasses, trees, and pineapples could illustrate the role of *SWN* in divergent species and the reason for its conservation. It is probable, considering the roles of *SWN* in *Arabidopsis* that the *SWN*-like genes play roles in embryogenesis, mega-gametogenesis or more generally in the haploid generation, for example in mosses. The J-A transition is often more defined in woody species like *Hedera helix* (Ivy), therefore investigating the *SWN*-like genes in woody species may prove more useful in deciphering its role in the vegetative transition.

5. 11. 5. Discovering the entirety of Pc-G function

Ultimately, it would be interesting to test the role of Pc-G function through ChIP chip and microarray analysis in a multitude specific cell types, developmental stages, and environmental conditions. This would allow the characterization of the entirety of Pc-G function in *Arabidopsis* ontogeny, morphogenesis, development, and response to external and internal influences.

The Pc-G is a complicated beast and the genomic scale tools currently being developed are ideally suited to further investigate the functions of the Pc-G, and may provide novel insight into the developmental workings of multi-cellular organisms. This knowledge in the future could be used to improve crop yields to alleviate the plight of hunger in the modern world.

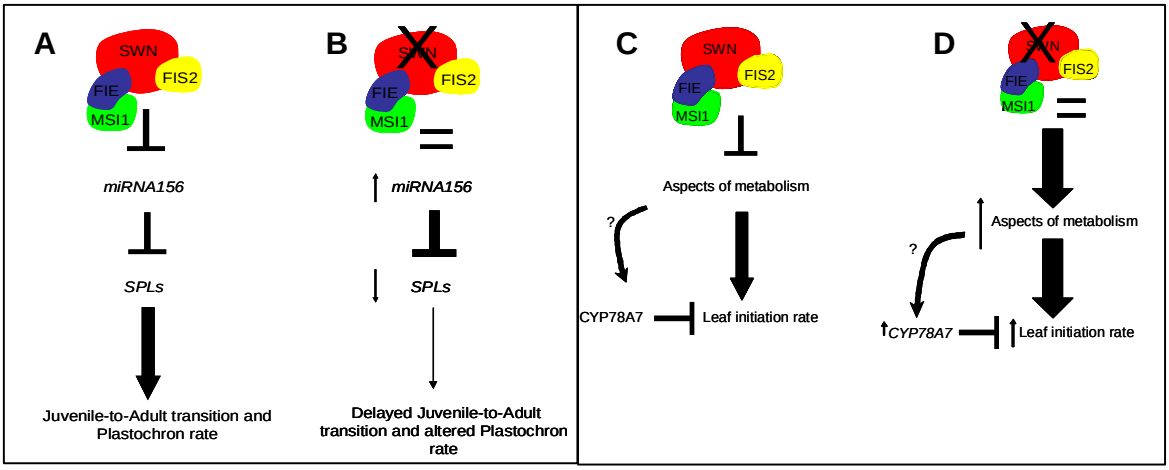

Fig. 5. 1. How *SWN* may regulate the Juvenile-Adult transition and leaf initiation rate.

A) In wild-type plants *SWN* acting in the Pc-G (independently of *CLF)*, or independently of the Pc-G may act to repress *miR156*, that causes the de-repression of *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) transcription factor clade. The expression of these genes induces adult leaf formation i.e. the J-A transition. B) In *swn*- plants *miRNA156* is not repressed, the increased expression of *miR156* ectopically represses the *SPLs* clade. This in turn causes the delayed J-A transition. C) In wild-type plants *SWN* acting in the Pc-G (independently of *CLF)* to potentially repress metabolism. The metabolites formed act as a signal to promote the plastochron rate, they may also act to promote *Cytochrome78A7*, which also acts to repress the metabolite signalling promoting leaf initiation rate. D) In *swn*- plants, there maybe a de-repression of the genes metabolising the J-A signal. The increased metabolite signal would cause increased plastochron rate, but the increase metabolites may increase the expression of *CYP78A7* that may act to buffer the effect.

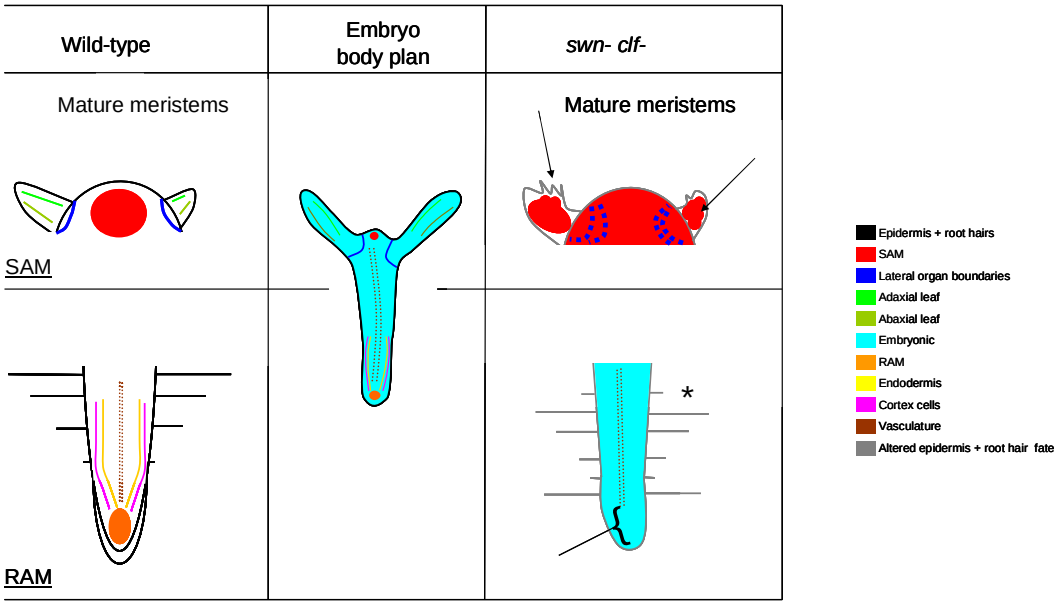

Fig. 5. 2. Schematic of meristem phenotypes in *swn- clf-* double mutants.
The embryonic body plan in both wild-type and *swn- clf-* , tissues include the epidermis, root and shoot meristems, vasculature, and leaf polarity. The wild-type plants maintain the differentiation pattern of SAM and RAM, leaves, roots, and epidermis, but there is a loss of embryo fate. *swn- clf-* mutants fail to maintain differentiation of the SAM and genes required for maintaining SAM identity expand into surrounding tissues. The lateral organ boundary is likely to be disrupted (indicated by the dotted line). The combination of SAM and lateral organ boundary fate is likely cause leaf serrations (arrows). The epidermis fate is disrupted in the aerial and below ground parts, which is like to cause the root hair initiation and morphology defects (*). The RAM is lost ({) and the root has acquired an embryonic fate.

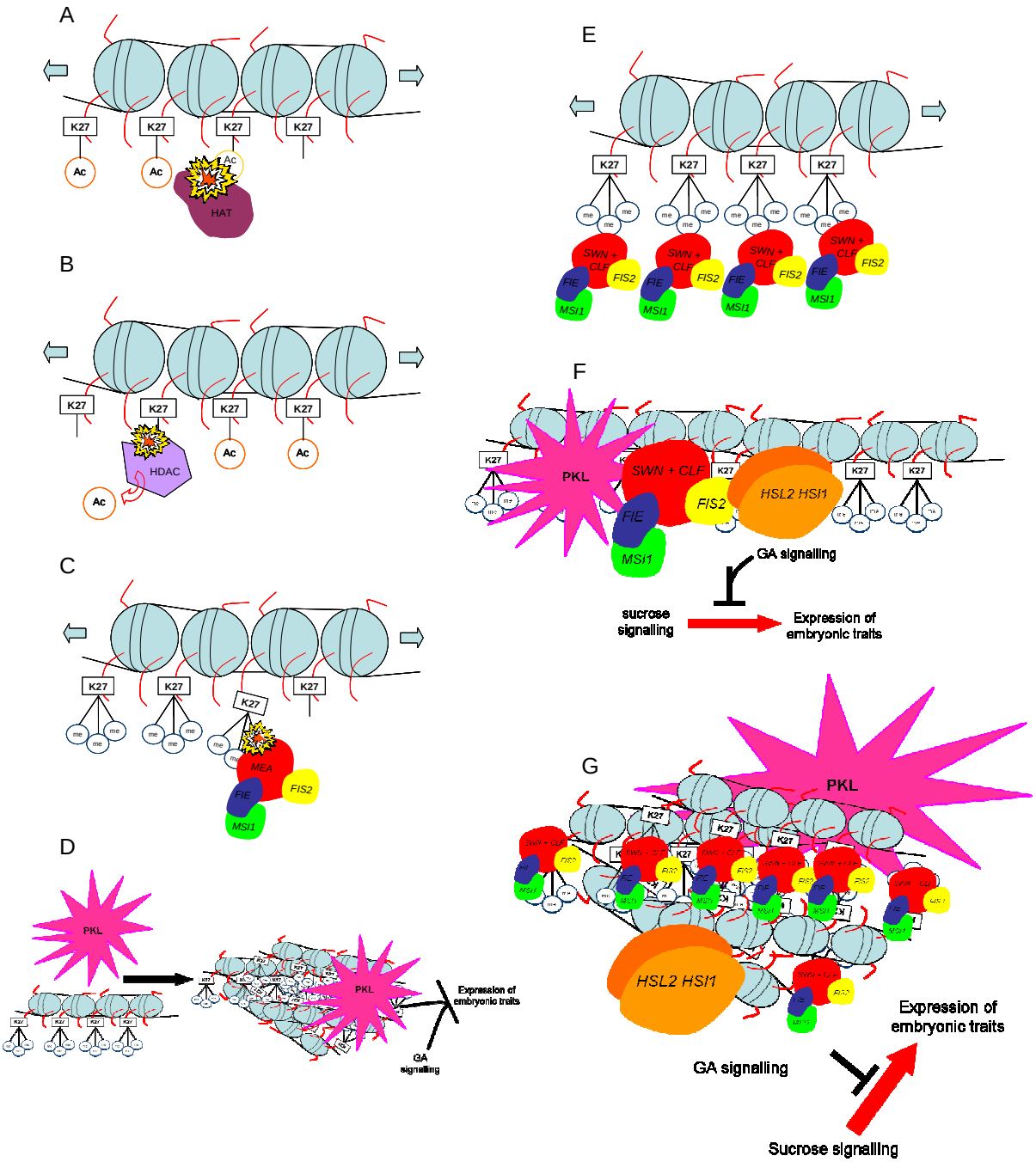
A
K27
Ac
HAT
B
HDAC
C
MEA
FIE
FIS2
MSI1
D
PKL
Expression of embryonic traits
GA signalling
E
SWN + CLF
FIS2
FIE
MSI1
F
PKL
HSL2 HSI1
GA signalling
sucrose signalling
Expression of embryonic traits
G
PKL
HSL2 HSI1
Expression of embryonic traits
GA signalling
Sucrose signalling

Fig. 5. 3. The proposed mechanism by which Pc-G, and other players, repress embryonic traits.
An example of how the Pc-G, *PKL*, and, *HSL1 HSI2* may repress *FUS3* and other embryonic trait master regulators from promoting and embryonic trait genes. (A) *FUS3* is activated by transcription factors and histone acetylation (perhaps of H3K27) during seed maturation. *FUS3* then activates down-stream targets, like *EM1* and other seed storage proteins. However, the MEA-Pc-G actively adds the H3K27 me^3 to the FUS3 loci to repress its expression in the endosperm (C). In late seed maturation *FUS3*, and probably the down-stream targets, are actively repressed by histone deacetylation (B), and the addition of H3K27me^3 created by the *SWN/ CLF*-Pc-G (E), the deacetylation may act as a signal for histone methylation, or be carried out by a putative member of the Pc-G. (D) Following histone methylation in seed maturation, the H3K27me^3 mark may be interpreted by *PKL*, and perhaps other *SWI/SNF* chromatin remodellers, to create localized regions of pseudo-heterochromatin (G). This causes the stable gene repression, in a method that GA signalling also inhibits their expression. *PKL* may or may not remain associated with the pseudo-heterochromatin. Post-germination, *PKL* and *SWN/CLF*-Pc-G act to maintain the H3K27me^3 mark and repressed state, perhaps by being enriched at the H3K27me^3 loci (E + F), or by being incorporated into the pseudo-heterochromatin (G). In addition to chromatin remodelling and histone methylation, *HSL2 HSI1* also act to repress sugar signalling to induce expression of embryonic traits, probably by inhibiting access of strong promoters or transcriptional machinery, seeming independently of chromatin state or histone modification (E + F). Light blue circles = nucleosomes, Red tails = histone tails, black = DNA, and K27 = histone 3 lysine 27. This mode of repression may be found at *FUS3* as indicated but probably occurs at the down-stream components of the embryonic trait pathway.

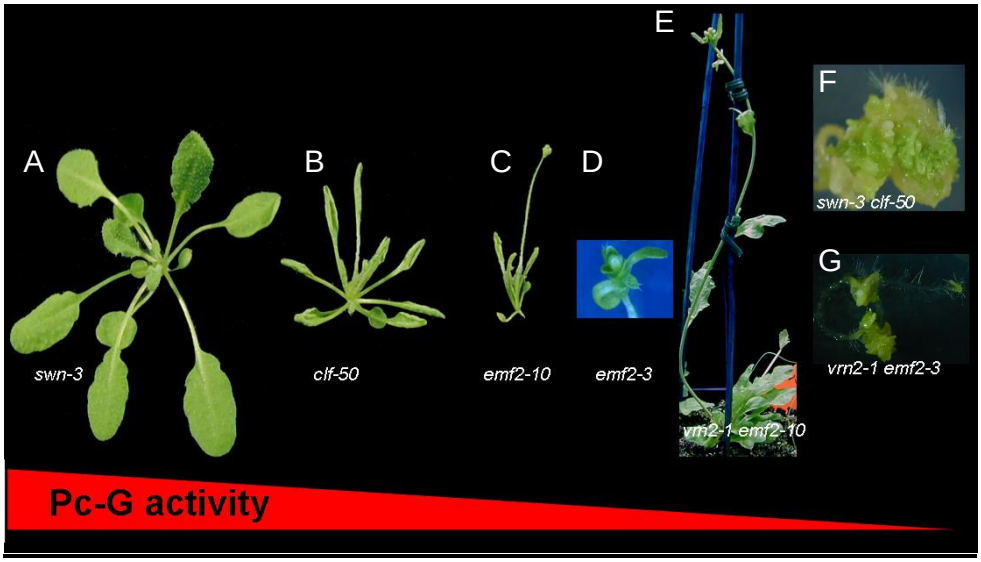

Fig. 5. 4. Gradient of Pc-G activity does not directly reflect Pc-G mutant phenotypes.
The phenotypes of Pc-G mutants, which often reflects the gradient of Pc-G function. Removing small amounts of Pc-G activity i.e. *SWN* (A), or *VRN* (whose phenotype is similar to *SWN)* has only subtle effects on development, whereas removing *CLF* (B) of *EMF2* (C + D) more severe defects observed indicating that they play an integral role in development. However, when *SWN* and *CLF* (F) or *VRN2 EMF2* (E + G) are mutated Pc-G activity is removed this can either result in a massive disruption in organogenesis which is masked by redundancy, or a suppressed phenotype. This indicates that the degree

of Pc-G activity often reflects the mutant Pc-G phenotype, but because of the Pc-G acting to repress antagonistic genes in the same pathway can result in unexpected phenotypes such as the *vrn2-1 emf2-10* double mutants (E).

6. 0. Materials and methods

6. 1. 0. Plant materials

The Columbia (Col), Wassilewskija (Ws), and Landsberg (Ler) accessions were used.

6. 1. 1. Mutant and transgenics lines

swn-1 is in a Ws background, and is described in Chanvivattana *et al.*, 2004. It contains A single insertion in 5'-UTR ~4 nucleotides from the ATG start codon. It is a weak allele as some full length transcript remains. *swn-2,-3,-7* are in Col background and were obtained from the Salk Institute Genome Analysis Laboratory collection. *swn-2* (Salk 010213) contains T-DNA inserts in introns, and *swn-3* and *-7* contain inserts in exons. *swn-3* (Salk 050195) and *swn-7* (SALK_109121) lines are prdicted to be null although this has not been proven, the other lines show reduced *SWN* expression but are more likely to be a partial loss of function. *clf-28* (SALK_139371) is in the Col background and was obtained form the Salk Institute Genome Analysis Laboratory collection, The T-DNA insertion in located in the 4^{th} exon and is a likely null mutation. *clf-81* is in the Col background and was isolated from an EMS mutagenesis screen, that carries a point mutation in the SET domain of the gene, suggesting that this is a null line (as described in Schubert *et al.*, 2006) and were provided by H. Tsukaya. Vernalization requiring Col ecotype plants carried a transgenic *FRI+* allele that originated from the H51 Stockholm ecotype, and seeds were provided by Prof. C. Dean (Johanson *et al.*, 2000). *emf2-3* allele is in a Col background and was found in an ethylmethanesulphonate (EMS) and has a small deletion as described in Sung *et al.*, 1992 and Yang *et al.*, 1995 . *vrn2-1* is in a mixed Ws and Ler background and was found in a mutagenesis screen (Gendall *et al.*, 2001; Sheldon *et al.*, 1999)

6. 1. 2. Marker lines

The SCR::GFP marker line used contained a transgene of the *SCR* promoter and gene fused to the Green Fluorescence Protein (GFP) and introduced in to the Col background,

donated by Dr. B. Scheres as described in Sabatini *et al.*, 2003. *SHR::GFP* marker line used contained *SHR* promoter and gene fused to the Green Fluorescence Protein (GFP) and introduced in to the Col background, and was also donated by Dr. B. Scheres and is described in Helariutta *et al.*, 2000. The Dr5::GUS marker line as described in main text is a synthetic auxin response marker, Dr5, which is comprised of 5 repeats of an auxin response factor binding sequence. The Dr5 element is fused to 4-methylumbelliferyl-β-glucuronidase (GUS) to identify the location of auxin activity. This transgene construct was introduced in to the Col background, and were donated by Dr. B. Scheres, and described in Ni *et al.,* 2001). Dr5::GFP marker line uses the same Dr5 repeat element but is fused to GFP, and introduced into the Col background. This line was donated by Dr. J. Friml, and is described in Friml *et al.,* 2003).

6. 2. Growth conditions for *Arabidopsis* on soil and sterile culture

For growing *Arabidopsis* on soil, the seeds where planted on 2:1:1mixture of Levington F2 compost, sand and fine grit, in pots of 120 mm x 80 mm. The seeds were planted out and stratified for 2-3 days in 4°C to encourage synchronized germination. No more than 8 plants per pot where allowed to grow. The seed trays were covered for ~1 week until seedlings had appeared. The conditions for growth were either long day (16 hours light and 8 hours dark) or short days (12 hours light and 12 hours dark) both at a temperature of 21-24°C.

For sterile culture seeds were sterilized, this was achieved by saturating the seeds in 70% ethanol 0.125% Nordent P40 solution for ~5 mins. This was followed by two washes for three minutes each in 95% ethanol. The seeds were then air dried in a sterile hood and plated out on Murashige-Skoog medium (MS) salts as growth media (1/2 x MS salts 0.7% agar 0.3% sucrose pH 5.7). For purpose of selection, antibiotic was added to the MS media; kanamycin was used at 50 μg/ml and hygromycin B was used at 40 μg/m. The seeds where stratified, and grown in the same conditions, as described above.

6. 3. Genetic Crosses

Using watch maker forceps, the sepals, petals and stamen were removed from a flower prior to anthesis leaving the carpel, the future female parent. Only closed flowers or floral buds were used ensuring that no self pollination had taken place and the most apical flower is used to reduce the chance of the "naked" carpel from being pollinated by anything other than the desired cross. The naked carpel was left to mature for 1-2 days. Following maturation of the carpel, the transfer of pollen was carried out. This was done by picking flowers from the intended male parent that were open and possessed mature pollen on the anther of the stamen, this was gently stroked over the now mature "naked" carpal until it is covered with the yellow pollen. After 2-3 weeks the silique was harvested before dehiscence.

6. 4. DNA extraction for PCR

A small amount of newly formed leaf (~5 mm^2) was harvested into a 1.5 ml micro-centrifuge tube. The tissue was quickly frozen in liquid nitrogen, after a few minutes the tissue was macerated using a small pestle. 400 µl of extraction buffer was added (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5 % SDS). This was vortexed and allowed to settle for 2 mins. Then the samples were centrifuged for 2 minutes at 1400 rpm in a microcentrifuge. 300 µl of the supernatant was placed in a clean 1.5 ml microcentrifuge tube and DNA was precipitated, and 300 µl of isopropanol was added. This was vortexed and allowed to settle for 2 minutes and then centrifuged at 14000 rpm for 2 minutes. The supernatant was tipped off and allowed to air dry. The DNA was then resuspended in 200 µl of 10 mM Tris pH 7.5. 2 µl of this was used in each 25µl PCR reaction.

6. 5. Polymerase Chain Reaction (PCR) for genotyping

To analyse the genotype of a plant from a family segregating T-DNA insertion alleles a sample of its DNA was extracted (as outlined above) and PCR genotyping was carried

out. I used duplex PCR reactions to discern the genotype of the plants, this involved three primers; one was a primer designed to anneal to the Left border of a T-DNA insertion, the other two were gene specific and flanked the T-DNA insertion. If no T-DNA insertion were present then only the product of the two gene specific primers would be detected at a known band size (or weight). If plants were homozygous for the T-DNA insertion then the left border primer and one of the gene specific primer products would be created, at a reduced size band. However, if a plant were heterozygous for the T-DNA insertion then both the gene specific product and the band from a gene specific primer and the left border primer would occur. The presence of PCR products were observed by running them out on an agarose gel, and observing the presence of DNA under Ultra Violet light (described below)

The PCR mix used is outlined below: they were prepared on ice in the thin-walled PCR tubes, then the PCR mixes were ran in a MJ Systems DNA engine thermocycler preheated to 94°C (so called “hot start”).

2 µl 10X magnesium-free PCR buffer (New England Biolabs (NEB) buffer)

2 µl $MgCl^2$ (10 mM)

0.4 µl dNTPs (10 mM)

0.2 µl TAQ polymerase (5 u/µl NEB)

0.8 µl Gene specific primer (10 µM)

0.8 µl Gene specific primer (10 µM)

0.8 µl Left border T-DNA primer (10 µM)

13.6 µl sterile water

1 µl DNA (approximately 50 ng)

A commonly used PCR program is outlined below:

Step 1:	94°C	2 mins
Step 2:	94°C	30 secs
Step 3:	60°C	30 secs
Step 4:	72°C	30 secs
Step 5:	72°C	10 mins
Step 6:	4°C	Forever

Steps from 2 to 4 were cycled 30 times.

The products from PCR reactions were then analysed by electrophoresis, usually, on a 1 % agarose gel, described below.

6. 6. Genotyping by dCAPs

In the case of *clf-81* it was not possible to use the triplex PCR as it is not a T-DNA insertion line, but the single nucleotide substitution creates a restriction site for the enodonulease restriction enzyme for *Hha*. 1. A PCR band is produced by two gene specific primers spanning the point mutation, then digested *Hha. 1*. The digest is then run on a gel electrophoresis to observe the number and sizes of bands to reveals its genotype. The PCR is exactly as above but two gene specific primers (see below) that spanned the point mutation, the volume is made up using water. The resultant PCR product was then digested with for one hour with *Hha*.1. The digest used 10 µl of the PCR reaction (~300 ng DNA) added to this is 2 µl buffer for the restriction enzyme (10 x) and the enzyme itself 0.5 µl. and incubated for one hour at 37°C, this was then ran gel electrophoresis 3% agarose gel. The genotype of the plants can be uncovered as the PCR spanned the point mutation, which created the restriction site, which when digested can

show one of three results 1) full length PCR band size, indicating that no restriction site was present and the DNA examined contains no DNA containing the point mutation, i.e. the plants are wild-type. 2) All the DNA shows a band size that has been digested, this illustrates all the DNA possessed the restriction site, and that the DNA came from a plant that was homozygous for the mutation. 3) Half of the bands are of the undigested size and half at the digested size, indicating that the plant from which the DNA came was heterozygous for the mutation.

6. 7. Agarose gel electrophoresis

Depending on the band sizes of interest 0.7 %, 1 % or 3 % agarose was weighed out and dissolved in 0.5 x TBE (89 mM Tris 89 mM Boric acid, 2 mM EDTA pH 8). This was then heated using a microwave, with frequent swilling, until the agarose had fully dissolved. It was then allowed to cool to approximately 60°C, at which point 5 µl of ethidium bromide (10 mg/ml stock) was added to every 100ml and mixed. This was then poured into a casting tray and left for 15 + minutes to set. The set agarose was then submerged in a gel tank in 0.5 x TBE. The PCR reactions (or dCAPS digests) were mixed with loading dye (0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 15 % ficoll (type 400)) to 1 % final concentration of loading dye to DNA sample. The samples were then added to the wells of the gel and underwent electrophoresis for ~1 hour at ~100V, depending on band sizes. When the bands on the gel were clearly resolved the bands were observed under a UV light and photographed. This was possible due to the addition of the ethidium bromide that interchelates between the base pairs of DNA and fluorescence under UV light, thus making the DNA bands visible. 6. 8. RNA isolation

RNA extractions, on samples described in the main body of text, were carried out using the Qiagen RNeasy plant mini kit, using the solutions provided and following the instruction therein.

6. 9. Reverse Transcription PCR (RT-PCR)

First strand cDNA was synthesised from RNA using Oligo dT primers which bind to the poly A tails that then underwent reverse transcription. This was carried out using the IMprom II™ reverse transcription kit, using the solutions provided and following the instruction therein.

The PCR reactions from the RT reactions were carried out as outlined above. To make the results semi-quantitative several PCR reactions were carried out with varying amounts dilutions of cDNA. One primer of the primer pair was designed with sequence homology to two contiguous exons (if this was possible). This was to discriminate cDNA from any possible genomic DNA contamination. *TUBULIN* (*TUB*) was used as an internal control analyzing relative amounts of cDNA present, as judged by the band fluorescence in UV light. Equal cDNA concentrations from each sample, was then used in the PCR reaction for the gene of interest. The intensity of the gene of interest PCR products were evaluated to find relative expression between samples.

6. 10. Quantitative RT-PCR

Q-RT-PCR reaction consisted of 12.5 μL of SYBR-green PCR mix (Sigma), 4 μL of cDNA, 7.5 μL of water, and 0.5 μL of each primer at 10 μM were used in triplicate 15 μL reactions. They where subjected to the following cycling conditions on Rotorgene RT3000 (Corbett Research, Australia): 50°C for 2 min, followed by 95°C for 10 min, then 40 cycles of 95°C for 15 s, and 60°C for 1 min. The SYBER-green mix contains a DNA polymerase and a fluorochrome. The fluorochrome specifically binds to double stranded DNA to emit fluorescence, which is captured by the Rotogene RT3000, the amount of fluorescence is directly correlated with the quantity of double stranded DNA. Therefore, each successive PCR cycle results in an increased DNA that bound by SYBR-green, therefore increasing intensity. Over time the amount of fluorescence and DNA quantity shows an “S” shaped curve, going through the lag, exponential, and plateau stages. Results where analyzed using Rotorgene (version 6) software. A single

amplification product in each reaction was confirmed by a dissociation curve and amplification quantified by comparison to a standard curve. Results for each tissue were normalized to the amplification of the *EIF*-4 control gene.

6. 11. Primers used for Genotyping

swn-7: Salk left border primer, 5′- GTCAGT TGTTCCACCTTGATCAC -3′ and 5′- GGTGCAGCTCTAGATTCTTTTG -3′. *clf-28:* Salk left border primer, 5′- CTGCCAGTTCAGGAATGGTT -3′ and 5′- GAAGGGAGCTCTCTGCTTGAT -3′. *clf-81*; 5′- GTTCTCGCGATCTATATCTTCGCG -3′ and 5′- GATGTTTCTGGTTGGGGAGCT -3′. *emf2-3;* Left border primer, 5′- AGTCAGAGAACATGGATGCGTATG -3′ and 5′- CGGACCGGGATAGTGAAGATGAAG -3′.

5 .11. 1. Primers used for RT-PCR and Q-RT-PCR

Locus	5′- > 3′	5′- > 3
EIF-4	TTCGCTCTTCTCTTTGCTCTC	GAACTCATCTTGTCCCTCAAGTA
CUC2	TCACAGTTGCTCCTCCTCCT	TCCAAGGATGAATGGGTGAT
ABI3	CAGGGATGGAAACCAGAAAA	TTCCAAACACGAGAGGTTCC
ABI4	CCACCGAACCAGCTAGAGAG	GATGGGACAATTCCAACACC
LBD40	CCTCCGTCCTGCGATATTTA	GCCAGTTTCCTGACCACAAT
LBD16	CCAACAACAGGTGGCTTTCT	GGTTGGTACTTTCCGAGCTG
AGL67	CAAGGATCGAGGACGTTTTG	CTGGACGAGGTTCGTTGATT
TT16	CAGGATGCCTCAACTCATTG	TGGAGGCTAAGTCATGTCCA
AIL5	TCCAGTCCATCTATGTCTTGTGA	GGAAACTACTCCGGTGGACA
AIL7	CTTTGCAACCGAAGAGGAAG	GAAGAGGACGGAAGCAACTG
EM1.	AGAACCCACAGGGACAACAG	GAAGGGATGTTGAAGGGACA
EXP. 2	TTCAGAAGGGTTCCATGTGAG	CACAACAGTCCGACCATCAC
GA-OX-2 (*AT5G07200*)	TTCCCGAAATCTTCGTATCC	GGCTTGTGAGAAGCAGAAGG
YUC4	ACATGTACAGAGTTTCTAACGACCAT	AACCGGTGTTTCCGAACATA
FLC	CGGTCTCATCGAGAAAGCTC	CCACAAGCTTGCTATCCACA
Similar to PRE1	ATGTCTTCTAGCAGAAGGTCGA	TGATACCGTTTGGAACGAC
Tubulin	GTTCTTGATAACGAGGCCTT	ACCTTCTTCCTCATCCTCG

5 .12. ATH1 microarray data analysis

RNA was extracted by myself as described above, and the NASC (Nottingham) carried out the cDNA synthesis, quality control, and the Affymetrix Genechip oligonucleotide ATH1 array (Affymetrix). Genechip arrays are synthetic DNA fragments (probes) placed on a quartz surface in specific locations. The provided cDNA samples are then labelled and hybridized to probes on the genechip, then the amount of label (fluorescence) can be detected for each probe set and this is directly correlated with quantity of cDNA, therefore expression level found in the plant tissue collected. The chip itself contains more than 22,500 probe-sets representing approximately 24,000 genes. Two biological replicates at 10 days old Col, *swn-7*, *clf-28*, and 12 day old *swn-7 clf-28* where used. The raw data from the Affymetrix service was imported into GENESPRING expression software. Background correction and normalization and gene expression analysis of the array data were performed using the GC-RMA routine in GeneSpring version 7.2 (Silicon Genetics), and absent signal were removed. Genes in the mutant backgrounds that displayed a two-fold difference in signal expression to wild-type where considered as mis-expressed. GENESPRING was used to assign mis-expressed genes Gene ontology (GO) terms then exported to Microsoft Excel. Gene lists were up loaded onto www.virtualplant.org for further comparisons and analysis of over-represented biological processes using the “Sungear” and “Biomaps” functions

5. 13. Propidium Iodine (PI) staining

PI stains cell walls and dead cells, and this was used to observe the internal architecture of wild-type and mutant seedling roots. This was carried out at varying ages by staining root tissues by mounting them on microscope slides using 10 mg/ml propidium iodine (PI), and placing a cover slip place on top and then photographed under the confocal microscope (described below). When PI is excited under 536 nm wave length light it emits a 620 nm wave length.

5. 14. GUS staining

To visualize gene expression patterns using the GUS reporter transgenes, the following steps were taken. Seedlings of various ages were placed in a six well plate, and the X-gluc (5-bromo-4-chloro-3-indoyl-beta- d- glucuronide) solution was added to the seedling to fully submerge them. The X-gluc, which consisted of 5 mM ferricyanide, 5mM ferrocyanide, 0.1 M NaH_2PO_4, 0.1 M Na_2HPO_4, 50 mM EDTA, 0.1 % Triton x100 and 0.25 mg/l X-gluc, pH 7. The seedlings in the solution were vacuum filtrated for 20 minutes and incubated overnight at 37°C. To remove the chlorophyll the seedlings were dehydrated through an ethanol series (30 %, 50 %, 70%, 80%) for 30 minutes each. and stored at room temperature in 70 % ethanol. A compound microscope and digital camera was used to take photographic images of the stained seedlings, they were placed in a 10cm petri dish, in liquid mounting solution (20 % ethanol, 20 % glycerol solution).

5. 15. Iodine potassium iodine (IKI) staining

Whole mounted seedlings were stained with IKI (5.7 mM iodine, 43.7 mM potassium iodine, 0.2M HCl) for 30 minutes at room temperature in 6 well plates. Then to remove the chlorophyll they were placed in 95 % ethanol for six hours and underwent choral hydrate clearing (see below).

5. 16. Sudan 7 staining

0.1 g of Sudan 7 was added to 10 ml of 60% polyethylene glycol (average molecular weight of 400) dissolved in water, this was heated 90°C for 1 hour to dissolve. An equal volume of 90% glycerol (in distilled water) was added. Samples were then submerged and incubated overnight. Subsequently the samples were rinsed several times with water, cleared using chloral hydrate (see below), and were photographed.

5. 17. Chloral hydrate clearing

Chloral hydrate was used to clear root tip stained with IKI or Sudan 7. This consisted of 1-2 minutes being submerged in chloral hydrate solution (8 g chloral hydrate, 2 ml water, 1 ml glycerol) and rinsed in water before the samples were analyzed by DIC microscopy. Embryos cleared were submerged in chloral hydrate, and photographed directly using DIC microscopy.

5. 18. Plant morphological and developmental analysis

The flowering time in long days was judged in two ways; Firstly, by counting how many leaves the plant possessed when a 1 cm bolt was formed. Secondly, it was recorded how many days it took for the 1 cm to be formed after germination. In short day conditions the same approaches were used, but leaf counting was carried out on a weekly basis to make sure no leaves were omitted. To find out how many juvenile leaves a plant possessed, it was recorded how many leaves were formed without abaxial trichomes until the first leaf with the abaxial trichomes (not including cotyledons), as leaves lacking abaxial trichomes are consider as juveniles, this was aided by using an dissecting microscope. The juvenile leaves are the first leaves that do not possess any abaxial trichomes not including the cotyledons.

5. 19. Microscopy

Kohler illumination is a means to optimize the light source and condenser set up for light field microscopy and was used on Nikon Eclipse E600 in DIC mode. Confocal microscopy was carried out using Olympus IX70 with an Olympus Fluroview driven light source. Light microscopy images were taken with a Nikon coolpix 950 digital camera. Images were trimmed and resized in Adode Photoshop 7.

SEM Microscopy was carried out using a Hitachi S4700II field emission scanning electron microscope using appropriate settings and was undertaken by Dr. C. Jeffry at the University of Edinburgh. Digital images were recorded using Hitachi FE PCSEM (version 3.2) software.

7. 0 References

Ahn O, Miyoshi K, Itoh JI, Nagato Y, Kurata N. 2002. A genetic and physical map of the region containing PLASTOCHRON1, a heterochronic gene, in rice (Oryza sativa L.). ***Theor.Appl.Genet.2002.Oct.;105.(5):654.-659.Epub.2002.Jun.14.*** **105: 654-659.**

Alvarez-Buylla ER, Garcia-Ponce B, Garay-Arroyo A. 2006. Unique and redundant functional domains of APETALA1 and CAULIFLOWER, two recently duplicated Arabidopsis thaliana floral MADS-box genes. ***J.Exp.Bot.2006.;57.(12.):3099.-107.Epub.2006.Aug.7.*** **57: 3099-3107.**

Amasino R. 2004. Take a cold flower. ***Nat.Genet.*** **36: 111-112.**

Amasino RM, Sheldon CC, Finnegan EJ, Rouse DT, Tadege M, Bagnall DJ, Helliwell CA, Peacock WJ, Dennis ES. 2005. Vernalization and flowering time The control of flowering by vernalization. ***Curr.Opin.Biotechnol.2005.Apr;16.(2):154.-8.*** **16: 154-158.**

Arroyo A, Bossi F, Finkelstein RR, Leon P. 2003. Three genes that affect sugar sensing (abscisic acid insensitive 4, abscisic acid insensitive 5, and constitutive triple response 1) are differentially regulated by glucose in Arabidopsis. ***Plant Physiol.2003.Sep.;133.(1):231.-42.*** **133: 231-242.**

Aubert D, Chen L, Moon YH, Martin D, Castle LA, Yang CH, Sung ZR. 2001. EMF1, a novel protein involved in the control of shoot architecture and flowering in ***Arabidopsis*****.** ***Plant Cell*** **13: 1865-1875.**

Ausio J, Dong F, van Holde KE. 1989. Use of selectively trypsinized nucleosome core particles to analyze the role of the histone "tails" in the stabilization of the nucleosome. ***J.Mol.Biol.*** **206: 451-463.**

Barkoulas M, Galinha C, Grigg SP, Tsiantis M. 2007. From genes to shape: regulatory interactions in leaf development. ***Curr.Opin.Plant Biol.2007.Dec.;10(6.):660.-6.Epub.2007.Sep.14.*** **10: 660-666.**

Bastow R, Dean C. 2003. Plant sciences. Deciding when to flower. ***Science.2003.Dec.5;302.(5651.):1695.-6.*** **302: 1695-1696.**

Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. ***Nature*** **427: 164-167.**

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G, Aalen RB. 2001. The Arabidopsis thaliana genome contains at

least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. ***Nucleic Acids Res.*** **29: 4319-4333.**

Ben Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. ***Nat.Genet.2008.May.;40.(5):499.-507.*** **40: 499-507.**

Beuchle D, Struhl G, Muller J. 2001. Polycomb group proteins and heritable silencing of Drosophila Hox genes. ***Development*** **128: 993-1004.**

Birve A, Sengupta AK, Beuchle D, Larsson J, Kennison JA, Rasmuson-Lestander A, Muller J. 2001. Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. ***Development*** **128: 3371-3379.**

Bishop G, Nomura T, Yokota T, Montoya T, Castle J, Harrison K, Kushiro T, Kamiya Y, Yamaguchi S, Bancos S, Szatmari AM, Szekeres M. 2006. Dwarfism and cytochrome P450-mediated C-6 oxidation of plant steroid hormones. ***Biochem.Soc.Trans.2006.Dec.;34.(Pt.6.):1199.-201.*** **34: 1199-1201.**

Blazquez MA, Soowal LN, Lee I, Weigel D. 1997. LEAFY expression and flower initiation in Arabidopsis. ***Development.*** **124: 3835-3844.**

Borner R, Kampmann G, Chandler J, Gleissner R, Wisman E, Apel K, Melzer S. 2000. A MADS domain gene involved in the transition to flowering in Arabidopsis. ***Plant J.2000.Dec.;24.(5):591.-9.*** **24: 591-599.**

Boss PK, Bastow RM, Mylne JS, Dean C. 2004. Multiple pathways in the decision to flower: enabling, promoting, and resetting. ***Plant Cell.2004.;16.Suppl:S18.-31.Epub.2004.Mar.22.*** **16 Suppl: S18-S31.**

Bouveret R, Schonrock N, Gruissem W, Hennig L. 2006. Regulation of flowering time by Arabidopsis MSI1. ***Development*** **133: 1693-1702.**

Bowen NJ, Fujita N, Kajita M, Wade PA. 2004. Mi-2/NuRD: multiple complexes for many purposes. ***Biochim.Biophys.Acta.2004.Mar.15.;1677.(1-3):52.-7.*** **1677: 52-57.**

Bowman JL, Smyth DR, Meyerowitz EM. 1989. Genes directing flower development in Arabidopsis. ***Plant Cell*** **1: 37-52.**

Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN. 2007. A high-resolution root spatiotemporal map reveals dominant expression patterns. ***Science.2007.Nov.2;318.(5851.):801.-6.*** **318: 801-806.**

Brand U, Grunewald M, Hobe M, Simon R. 2002. Regulation of CLV3 expression by two homeobox genes in Arabidopsis. ***Plant Physiol.2002.Jun.;129.(2):565.-75.*** **129: 565-575.**

Breiling A, Orlando V. 2002. SET domain proteins reSET gene expression. ***Nat.Struct.Biol.*** **9: 894-896.**

Calonje M, Sanchez R, Chen L, Sung ZR. 2008. EMBRYONIC FLOWER1 Participates in Polycomb Group-Mediated AG Gene Silencing in Arabidopsis. ***Plant Cell.2008.Feb.;20.(2):277.-91.Epub.2008.Feb.15.*** **20: 277-291.**

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. ***Science.2002.Nov.1;298.(5595.):1039.-43.Epub.2002.Sep.26.*** **298: 1039-1043.**

Cao R, Zhang Y. 2004. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. ***Curr.Opin.Genet.Dev.2004.Apr;14(2):155.-64.*** **14: 155-164.**

Carles C, Bies-Etheve N, Aspart L, Leon-Kloosterziel KM, Koornneef M, Echeverria M, Delseny M. 2002. Regulation of Arabidopsis thaliana Em genes: role of ABI5. ***Plant J.2002.May.;30.(3):373.-83.*** **30: 373-383.**

Chamberlain SJ, Yee D, Magnuson T. 2008. Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. ***Stem Cells.2008.Jun.;26.(6.):1496.-505.Epub.2008.Apr 10.*** **26: 1496-1505.**

Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J. 2004. Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. ***Development*** **131: 5263-5276.**

Chassot C, Buchala A, Schoonbeek HJ, Metraux JP, Lamotte O. 2008. Wounding of Arabidopsis leaves causes a powerful but transient protection against Botrytis infection. ***Plant J.2008.Apr 30.;.***

Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ. 1997. Fertilization-independent seed development in Arabidopsis thaliana. ***Proc.Natl.Acad.Sci.U.S.A*** **94: 4223-4228.**

Cheah KS, Osborne DJ. 1977. Analysis of nucleosomal deoxyribonucleic acid in a higher plant. ***Biochem.J.*** **163: 141-144.**

Chen L, Cheng JC, Castle L, Sung ZR. 1997. EMF genes regulate Arabidopsis inflorescence development. ***Plant Cell*** **9: 2011-2024.**

Cheng Y, Dai X, Zhao Y. 2007. Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in Arabidopsis. *Plant Cell.2007.Aug.;19.(8.):2430.-9.Epub.2007.Aug.17.* 19: 2430-2439.

Clarke JH, Tack D, Findlay K, Van Montagu M, Van Lijsebettens M. 1999. The SERRATE locus controls the formation of the early juvenile leaves and phase length in Arabidopsis. *Plant J.* 20: 493-501.

Coen ES, Meyerowitz EM. 1991. The war of the whorls: genetic interactions controlling flower development. *Nature.* 353: 31-37.

Costa S, Shaw P. 2006. Chromatin organization and cell fate switch respond to positional information in Arabidopsis. *Nature.2006.Jan.26.;439.(7075.):493.-6.Epub.2005.Dec.18.* 439: 493-496.

Crevillen, P. & Dean, C. 2008. "Dynamic association of different components of the vernalization specific-Polycomb complex determines the epigenetic regulation of the floral repressor FLC". Conference of Chromatin at the nexus of cell division and differentiation, Madrid, Jun.

Cui H, Levesque MP, Vernoux T, Jung JW, Paquette AJ, Gallagher KL, Wang JY, Blilou I, Scheres B, Benfey PN. 2007. An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. *Science.2007.Apr 20.;316.(5823.):421.-5.* 316: 421-425.

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. 2002. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111: 185-196.

Davies B, Motte P, Keck E, Saedler H, Sommer H, Schwarz-Sommer Z. 1999. PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. *EMBO J.* 18: 4023-4034.

de la Cruz CC, Kirmizis A, Simon MD, Isono K, Koseki H, Panning B. 2007. The polycomb group protein SUZ12 regulates histone H3 lysine 9 methylation and HP1 alpha distribution. *Chromosome.Res.2007.;15.(3):299.-314.Epub.2007.May.10.* 15: 299-314.

Dean Rider S Jr, Henderson JT, Jerome RE, Edenberg HJ, Romero-Severson J, Ogas J. 2003. Coordinate repression of regulators of embryonic identity by PICKLE during germination in Arabidopsis. *Plant J.2003.Jul.;35.(1):33.-43.* 35: 33-43.

Debeaujon I, Koornneef M. 2000. Gibberellin requirement for Arabidopsis seed germination is determined both by testa characteristics and embryonic abscisic acid. ***Plant Physiol.2000.Feb.;122.(2):415.-24.*** **122: 415-424.**

Dejardin J, Cavalli G. 2005. Epigenetic inheritance of chromatin states mediated by Polycomb and trithorax group proteins in Drosophila. ***Prog.Mol.Subcell.Biol.*** **38: 31-63.**

Di Laurenzio L, Wysocka-Diller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, Hahn MG, Feldmann KA, Benfey PN. 1996. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. ***Cell.*** **86: 423-433.**

Dolan L, Janmaat K, Willemsen V, Linstead P, Poethig S, Roberts K, Scheres B. 1993. Cellular organisation of the Arabidopsis thaliana root. ***Development.*** **119: 71-84.**

Elgin SC, Grewal SI. 2003. Heterochromatin: silence is golden. ***Curr.Biol.2003.Dec.2;13(23):R895.-8.*** **13: R895-R898.**

Eriksson S, Bohlenius H, Moritz T, Nilsson O. 2006. GA4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. ***Plant Cell.2006.Sep.;18.(9.):2172.-81.Epub.2006.Aug.18.*** **18: 2172-2181.**

Exner V, Hennig L. 2008. Chromatin rearrangements in development. ***Curr.Opin.Plant Biol.2008.Feb.;11(1):64.-9.Epub.2007.Nov.*** **11: 64-69.**

Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, Carrington JC. 2006. Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. ***Curr.Biol.2006.May.9.;16.(9.):939.-44.*** **16: 939-944.**

Farkas G, Gausz J, Galloni M, Reuter G, Gyurkovics H, Karch F. 1994. The Trithorax-like gene encodes the Drosophila GAGA factor. ***Nature.*** **371: 806-808.**

Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, Ditta G, Yanofsky MF, Kater MM, Colombo L. 2003. MADS-box protein complexes control carpel and ovule development in Arabidopsis. ***Plant Cell*** **15: 2603-2611.**

Francis NJ, Kingston RE. 2001. Mechanisms of transcriptional memory. ***Nat.Rev.Mol.Cell Biol.*** **2: 409-421.**

Friml J, Vieten A, Sauer M, Weijers D, Schwarz H, Hamann T, Offringa R, Jurgens G. 2003. Efflux-dependent auxin gradients establish the apical-basal axis of Arabidopsis. ***Nature.2003.Nov.13;426.(6963.):147.-53.*** **426: 147-153.**

Gallois JL, Woodward C, Reddy GV, Sablowski R. 2002. Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. ***Development.2002.Jul.;129.(13):3207.-17.*** **129: 3207-3217.**

Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, Saedler H, Huijser P. 2007. The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. ***Plant J.2007.Feb.;49.(4):683.-93.Epub.2007.Jan.8.*** **49: 683-693.**

Garces HM, Champagne CE, Townsley BT, Park S, Malho R, Pedroso MC, Harada JJ, Sinha NR. 2007. Evolution of asexual reproduction in leaves of the genus Kalanchoe. ***Proc.Natl.Acad.Sci.U.S.A.2007.Sep.25.;104.(39.):15578.-83.Epub.2007.Sep.24.*** **104: 15578-15583.**

Gendall AR, Levy YY, Wilson A, Dean C. 2001. The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis. ***Cell*** **107: 525-535.**

Glawischnig E. 2006. The role of cytochrome P450 enzymes in the biosynthesis of camalexin. ***Biochem.Soc.Trans.2006.Dec.;34.(Pt.6.):1206.-8.*** **34: 1206-1208.**

Glicksman MA, Brower DL. 1988. Misregulation of homeotic gene expression in Drosophila larvae resulting from mutations at the extra sex combs locus. ***Dev.Biol.*** **126: 219-227.**

Gomez-Mena C, de Folter S, Costa MM, Angenent GC, Sablowski R. 2005. Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. ***Development*** **132: 429-438.**

Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G. 1997. A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. ***Nature*** **386: 44-51.**

Goodrich J, Tweedie S. 2002. REMEMBRANCE OF THINGS PAST: Chromatin Remodeling in Plant Development. ***Annu.Rev.Cell Dev.Biol.*** **18: 707-746.**

Gray WM. 2004. Hormonal regulation of plant growth and development. ***PLoS.Biol.2004.Sep.;2(9.):E311.Epub.2004.Sep.14.*** **2: E311.**

Grieneisen VA, Xu J, Maree AF, Hogeweg P, Scheres B. 2007. Auxin transport is sufficient to generate a maximum and gradient guiding root growth. ***Nature.2007.Oct.25.;449.(7165.):1008.-13.*** **449: 1008-1013.**

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB. 1998. Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. ***Science*** **280: 446-450.**

Gubler F, Millar AA, Jacobsen JV. 2005. Dormancy release, ABA and pre-harvest sprouting. ***Curr.Opin.Plant Biol.2005.Apr;8.(2):183.-7.*** **8: 183-187.**

Guitton AE, Page DR, Chambrier P, Lionnet C, Faure JE, Grossniklaus U, Berger F. 2004. Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis thaliana. ***Development*** **131: 2971-2981.**

Haecker A, Gross-Hardt R, Geiges B, Sarkar A, Breuninger H, Herrmann M, Laux T. 2004. Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in Arabidopsis thaliana. ***Development.2004.Feb.;131.(3):657.-68.Epub.2004.Jan.7.*** **131: 657-668.**

Halim VA, Vess A, Scheel D, Rosahl S. 2006. The role of salicylic acid and jasmonic acid in pathogen defence. ***Plant Biol.(Stuttg).2006.May.;8.(3):307.-13.*** **8: 307-313.**

Hamberger B, Bohlmann J. 2006. Cytochrome P450 mono-oxygenases in conifer genomes: discovery of members of the terpenoid oxygenase superfamily in spruce and pine. ***Biochem.Soc.Trans.2006.Dec.;34.(Pt.6.):1209.-14.*** **34: 1209-1214.**

Hannah-Alava A. 1958. Developmental Genetics of the Posterior Legs in Drosophila Melanogaster. ***Genetics.*** **43: 878-905.**

Harding K, Levine M. 1988. Gap genes define the limits of antennapedia and bithorax gene expression during early development in Drosophila. ***EMBO J.*** **7: 205-214.**

Hartig K, Beck E. 2006. Crosstalk between auxin, cytokinins, and sugars in the plant cell cycle. ***Plant Biol.(Stuttg).2006.May.;8.(3):389.-96.*** **8: 389-396.**

Hay A, Tsiantis M. 2006. The genetic basis for differences in leaf form between Arabidopsis thaliana and its wild relative Cardamine hirsuta. ***Nat.Genet.2006.Aug.;38.(8.):942.-7.Epub.2006.Jul.2.*** **38: 942-947.**

Hayama R, Coupland G. 2003. Shedding light on the circadian clock and the photoperiodic control of flowering. ***Curr.Opin.Plant Biol.*** **6: 13-19.**

Heidstra R, Welch D, Scheres B. 2004. Mosaic analyses using marked activation and deletion clones dissect Arabidopsis SCARECROW action in asymmetric cell division. ***Genes Dev.*** **18: 1964-1969.**

Helariutta Y, Fukaki H, Wysocka-Diller J, Nakajima K, Jung J, Sena G, Hauser MT, Benfey PN. 2000. The SHORT-ROOT gene controls radial patterning of the Arabidopsis root through radial signaling. ***Cell*** **101: 555-567.**

Helliwell CA, Chin-Atkins AN, Wilson IW, Chapple R, Dennis ES, Chaudhury A. 2001. The Arabidopsis AMP1 gene encodes a putative glutamate carboxypeptidase. *Plant Cell.2001.Sep.;13(9.):2115.-25.* 13: 2115-2125.

Henderson JT, Li HC, Rider SD, Mordhorst AP, Romero-Severson J, Cheng JC, Robey J, Sung ZR, de Vries SC, Ogas J. 2004. PICKLE acts throughout the plant to repress expression of embryonic traits and may play a role in gibberellin-dependent responses. *Plant Physiol.2004.Mar.;134.(3):995.-1005.Epub.2004.Feb.12.* 134: 995-1005.

Hennig L, Taranto P, Walser M, Schonrock N, Gruissem W. 2003. Arabidopsis MSI1 is required for epigenetic maintenance of reproductive development. *Development* 130: 2555-2565.

Hsieh TF, Fischer RL. 2005. Biology of chromatin dynamics. *Annu.Rev.Plant Biol.2005.;56.:327.-51.* 56: 327-351.

Hunter C, Willmann MR, Wu G, Yoshikawa M, de la Luz Gutierrez-Nava, Poethig SR. 2006. Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development.2006.Aug.;133.(15.):2973.-81.Epub.2006.Jul.3.* 133: 2973-2981.

Hunter CA, Aukerman MJ, Sun H, Fokina M, Poethig RS. 2003. PAUSED encodes the Arabidopsis exportin-t ortholog. *Plant Physiol.2003.Aug.;132.(4):2135.-43.* 132: 2135-2143.

Ikeda-Iwai M, Satoh S, Kamada H. 2002. Establishment of a reproducible tissue culture system for the induction of Arabidopsis somatic embryos. *J.Exp.Bot.2002.Jul.;53.(374.):1575.-80.* 53: 1575-1580.

Ikura T, Tashiro S, Kakino A, Shima H, Jacob N, Amunugama R, Yoder K, Izumi S, Kuraoka I, Tanaka K, Kimura H, Ikura M, Nishikubo S, Ito T, Muto A, Miyagawa K, Takeda S, Fishel R, Igarashi K, Kamiya K. 2007. DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics. *Mol.Cell Biol.2007.Oct.;27.(20.):7028.-40.Epub.2007.Aug.20.* 27: 7028-7040.

Imaizumi T, Kay SA. 2006. Photoperiodic control of flowering: not only by coincidence. *Trends Plant Sci.2006.Nov.;11(11):550.-8.Epub.2006.Oct.10.* 11: 550-558.

Jack T, Fox GL, Meyerowitz EM. 1994. Arabidopsis homeotic gene APETALA3 ectopic expression: transcriptional and posttranscriptional regulation determine floral organ identity. *Cell* 76: 703-716.

Jackson JP, Lindroth AM, Cao X, Jacobsen SE. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416: 556-560.

Jacobs SA, Khorasanizadeh S. 2002. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295: 2080-2083.

Jenik PD, Barton MK. 2005. Surge and destroy: the role of auxin in plant embryogenesis. *Development.2005.Aug.;132.(16.):3577.-85.* 132: 3577-3585.

Jenik PD, Gillmor CS, Lukowitz W. 2007. Embryonic patterning in Arabidopsis thaliana. *Annu.Rev.Cell Dev.Biol.2007.;23:207.-36.* 23: 207-236.

Jenuwein T, Allis CD. 2001. Translating the histone code. *Science* 293: 1074-1080.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C. 2000. Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science* 290: 344-347.

Johnston AJ, Meier P, Gheyselinck J, Wuest SE, Federer M, Schlagenhauf E, Becker JD, Grossniklaus U. 2007. Genetic subtraction profiling identifies genes essential for Arabidopsis reproduction and reveals interaction between the female gametophyte and the maternal sporophyte. *Genome Biol.2007.;8.(10):R204.* 8: R204.

Jones RS, Gelbart WM. 1990. Genetic analysis of the enhancer of zeste locus and its role in gene regulation in Drosophila melanogaster. *Genetics* 126: 185-199.

Jullien PE, Katz A, Oliva M, Ohad N, Berger F. 2006. Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. *Curr.Biol.* 16: 486-492.

Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A, Hattori T. 2005. LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant Cell Physiol.2005.Mar.;46.(3):399.-406.Epub.2005.Feb.2.* 46: 399-406.

Kahn TG, Schwartz YB, Dellino GI, Pirrotta V. 2006. Polycomb complexes and the propagation of the methylation mark at the Drosophila ubx gene. *J.Biol.Chem.2006.Sep.29.;281.(39.):29064.-75.Epub.2006.Aug.2.* 281: 29064-29075.

Kanno R, Janakiraman H, Kanno M. 2008. Epigenetic regulator polycomb group protein complexes control cell fate and cancer. *Cancer Sci.2008.Jun.;99.(6.):1077.-84.Epub.2008.Apr 14.* 99: 1077-1084.

Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D. 1999. Activation tagging of the floral inducer FT. *Science.* 286: 1962-1965.

Katz A, Oliva M, Mosquna A, Hakim O, Ohad N. 2004. FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. *Plant J.* 37: 707-719.

Kempin SA, Savidge B, Yanofsky MF. 1995. Molecular basis of the cauliflower phenotype in Arabidopsis. *Science.* 267: 522-525.

Kim GT, Tsukaya H, Uchimiya H. 1998. The CURLY LEAF gene controls both division and elongation of cells during the expansion of the leaf blade in Arabidopsis thaliana. *Planta* 206: 175-183.

Kim JI, Sharkhuu A, Jin JB, Li P, Jeong JC, Baek D, Lee SY, Blakeslee JJ, Murphy AS, Bohnert HJ, Hasegawa PM, Yun DJ, Bressan RA. 2007. yucca6, a dominant mutation in Arabidopsis, affects auxin accumulation and auxin-related phenotypes. *Plant Physiol.2007.Nov.;145.(3):722.-35.Epub.2007.Sep.20.* 145: 722-735.

Kinoshita T, Harada JJ, Goldberg RB, Fischer RL. 2001. Polycomb repression of flowering during early plant development. *Proc.Natl.Acad.Sci.U.S.A* 98: 14156-14161.

Kiss JZ, Hertel R, Sack FD. 1989. Amyloplasts are necessary for full gravitropic sensitivity in roots of Arabidopsis thaliana. *Planta.* 177: 198-206.

Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, Katz A, Margossian L, Harada JJ, Goldberg RB, Fischer RL. 1999. Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc.Natl.Acad.Sci.U.S.A* 96: 4186-4191.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science.* 286: 1960-1962.

Kohler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W. 2003a. Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J.* 22: 4804-4814.

Kohler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U. 2003b. The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev.* 17: 1540-1553.

Kohler C, Page DR, Gagliardini V, Grossniklaus U. 2005. The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. *Nat.Genet.* 37: 28-30.

Komeda Y. 2004. Genetic regulation of time to flower in Arabidopsis thaliana. *Annu.Rev.Plant Biol.2004.;55.:521.-35.* 55: 521-535.

Koornneef M, Hanhart CJ, Hilhorst HW, Karssen CM. 1989. In Vivo Inhibition of Seed Development and Reserve Protein Accumulation in Recombinants of Abscisic Acid Biosynthesis and Responsiveness Mutants in Arabidopsis thaliana. *PLANT PHYSIOLOGY* 90: 463-469.

Kotake T, Takada S, Nakahigashi K, Ohto M, Goto K. 2003. Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. *Plant Cell Physiol.2003.Jun.;44.(6.):555.-64.* 44: 555-564.

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410: 116-120.

Lee H, Suh SS, Park E, Cho E, Ahn JH, Kim SG, Lee JS, Kwon YM, Lee I. 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes Dev.2000.Sep.15.;14(18.):2366.-76.* 14: 2366-2376.

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA. 2006. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125: 301-313.

Lenhard M, Jurgens G, Laux T. 2002. The WUSCHEL and SHOOTMERISTEMLESS genes fulfil complementary roles in Arabidopsis shoot meristem regulation. *Development.2002.Jul.;129.(13):3195.-206.* 129: 3195-3206.

Leyser O. 2005. Auxin distribution and plant pattern formation: how many angels can dance on the point of PIN? *Cell.2005.Jun.17;121.(6.):819.-22.* 121: 819-822.

Li H, He Z, Lu G, Lee SC, Alonso J, Ecker JR, Luan S. 2007. A WD40 domain cyclophilin interacts with histone H3 and functions in gene repression and organogenesis in Arabidopsis. *Plant Cell.2007.Aug.;19.(8.):2403.-16.Epub.2007.Aug.17.* 19: 2403-2416.

Li HC, Chuang K, Henderson JT, Rider SD, Jr., Bai Y, Zhang H, Fountain M, Gerber J, Ogas J. 2005. PICKLE acts during germination to repress expression of embryonic traits. *Plant J.2005.Dec.;44.(6.):1010.-22.* 44: 1010-1022.

Liljegren SJ, Gustafson-Brown C, Pinyopich A, Ditta GS, Yanofsky MF. 1999. Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 specify meristem fate. *Plant Cell.* 11: 1007-1018.

Lim PO, Kim HJ, Nam HG. 2007. Leaf senescence. *Annu.Rev.Plant Biol.2007.;58.:115.-36.* 58: 115-136.

Lin MS, Alfi OS. 1976. Detection of sister chromatid exchanges by 4'-6-diamidino-2-phenylindole fluorescence. *Chromosoma.* 57: 219-225.

Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I, Jenuwein T, Khorasanizadeh S, Jacobsen SE. 2004. Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *EMBO J.* 23: 4286-4296.

Lohmann JU, Hong RL, Hobe M, Busch MA, Parcy F, Simon R, Weigel D. 2001. A molecular link between stem cell regulation and floral patterning in Arabidopsis. *Cell* 105: 793-803.

Long JA, Moan EI, Medford JI, Barton MK. 1996. A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. *Nature* 379: 66-69.

Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH. 2002. ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J.2002.Nov.;32.(3):317.-28.* 32: 317-328.

Luerssen H, Kirik V, Herrmann P, Misera S. 1998. FUSCA3 encodes a protein with a conserved VP1/AB13-like B3 domain which is of functional importance for the regulation of seed maturation in Arabidopsis thaliana. *Plant J.* 15: 755-764.

Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudhury A. 2000. Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proc.Natl.Acad.Sci.U.S.A* 97: 10637-10642.

Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM. 1999. Genes controlling fertilization-independent seed development in Arabidopsis thaliana. *Proc.Natl.Acad.Sci.U.S.A* 96: 296-301.

Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, Grossniklaus U, Kohler C. 2006. Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep.* 7: 947-952.

Martienssen RA, Colot V. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science* 293: 1070-1074.

Matzke MA, Matzke AJ. 1998. Epigenetic silencing of plant transgenes as a consequence of diverse cellular defence responses. *Cell Mol.Life Sci.* 54: 94-103.

Mayama T, Ohtsubo E, Tsuchimoto S. 2003. Isolation and expression analysis of petunia CURLY LEAF-like genes. *Plant Cell Physiol.2003.Aug.;44.(8.):811.-9.* 44: 811-819.

Mayer U, Buttner G, Jurgens G. 1993. Apical-basal pattern formation in the Arabidopsis embryo: studies on the role of the gnom gene. *Development* 117: 149-162.

Metsuyanim S, Pode-Shakked N, Schmidt-Ott KM, Keshet G, Rechavi G, Blumental D, Dekel B. 2008. Accumulation of malignant renal stem cells is associated with epigenetic changes in normal renal progenitor genes. *Stem Cells.2008.Jul.;26.(7.):1808.-17.Epub.2008.May.8.* 26: 1808-1817.

Michaels SD, Amasino RM. 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11: 949-956.

Miyoshi K, Ahn BO, Kawakatsu T, Ito Y, Itoh J, Nagato Y, Kurata N. 2004. PLASTOCHRON1, a timekeeper of leaf initiation in rice, encodes cytochrome P450. *Proc.Natl.Acad.Sci.U.S.A.2004.Jan.20.;101.(3):875.-80.Epub.2004.Jan.7.* 101: 875-880.

Mizukami Y, Ma H. 1992. Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. *Cell.* 71: 119-131.

Mohd-Sarip A, van der Knaap JA, Wyman C, Kanaar R, Schedl P, Verrijzer CP. 2006. Architecture of a polycomb nucleoprotein complex. *Mol.Cell.2006.Oct.6.;24.(1):91.-100.* 24: 91-100.

Moon YH, Chen L, Pan RL, Chang HS, Zhu T, Maffeo DM, Sung ZR. 2003b. EMF genes maintain vegetative development by repressing the flower program in Arabidopsis. *Plant Cell* 15: 681-693.

Moon YH, Chen L, Pan RL, Chang HS, Zhu T, Maffeo DM, Sung ZR. 2003a. EMF genes maintain vegetative development by repressing the flower program in Arabidopsis. *Plant Cell* 15: 681-693.

Mylne JS, Barrett L, Tessadori F, Mesnage S, Johnson L, Bernatavichute YV, Jacobsen SE, Fransz P, Dean C. 2006. LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. *Proc.Natl.Acad.Sci.U.S.A* 103: 5012-5017.

Mylona P, Linstead P, Martienssen R, Dolan L. 2002. SCHIZORIZA controls an asymmetric cell division and restricts epidermal identity in the Arabidopsis root. ***Development.2002.Sep.;129.(18.):4327.-34.*** **129: 4327-4334.**

Nakahigashi K, Jasencakova Z, Schubert I, Goto K. 2005. The Arabidopsis heterochromatin protein1 homolog (TERMINAL FLOWER2) silences genes within the euchromatic region but not genes positioned in heterochromatin. ***Plant Cell Physiol.2005.Nov.;46.(11):1747.-56.Epub.2005.Aug.29.*** **46: 1747-1756.**

Nakajima K, Sena G, Nawy T, Benfey PN. 2001a. Intercellular movement of the putative transcription factor SHR in root patterning. ***Nature.2001.Sep.20.;413.(6853.):307.-11.*** **413: 307-311.**

Nakajima K, Sena G, Nawy T, Benfey PN. 2001b. Intercellular movement of the putative transcription factor SHR in root patterning. ***Nature*** **413: 307-311.**

Nemhauser JL, Zambryski PC, Roe JL. 1998. Auxin signaling in Arabidopsis flower development? ***Curr.Opin.Plant Biol.*** **1: 531-535.**

Ng J, Hart CM, Morgan K, Simon JA. 2000. A Drosophila ESC-E(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC. ***Mol.Cell Biol.*** **20: 3069-3078.**

Nowack MK, Shirzadi R, Dissmeyer N, Dolf A, Endl E, Grini PE, Schnittger A. 2007. Bypassing genomic imprinting allows seed development. ***Nature.2007.May.17;447.(7142.):312.-5.Epub.2007.Apr 29.*** **447: 312-315.**

Ogas J, Cheng JC, Sung ZR, Somerville C. 1997. Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. ***Science.*** **277: 91-94.**

Ogas J, Kaufmann S, Henderson J, Somerville C. 1999. PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. ***Proc.Natl.Acad.Sci.U.S.A*** **96: 13839-13844.**

Ohad N, Margossian L, Hsu YC, Williams C, Repetti P, Fischer RL. 1996. A mutation that allows endosperm development without fertilization. ***Proc.Natl.Acad.Sci.U.S.A*** **93: 5319-5324.**

Ohad N, Yadegari R, Margossian L, Hannon M, Michaeli D, Harada JJ, Goldberg RB, Fischer RL. 1999. Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. ***Plant Cell*** **11: 407-416.**

Okamuro JK, den Boer BG, Lotys-Prass C, Szeto W, Jofuku KD. 1996. Flowers into shoots: photo and hormonal control of a meristem identity switch in Arabidopsis. ***Proc.Natl.Acad.Sci.U.S.A*** **93: 13831-13836.**

Olins AL, Olins DE. 1974. Spheroid chromatin units (v bodies). ***Science*** **183: 330-332.**

Papait R, Pistore C, Grazini U, Babbio F, Cogliati S, Pecoraro D, Brino L, Morand AL, Dechampesme AM, Spada F, Leonhardt H, McBlane F, Oudet P, Bonapace IM. 2008. The PHD Domain of Np95 (mUHRF1) Is Involved in Large-Scale Reorganization of Pericentromeric Heterochromatin. ***Mol.Biol.Cell.2008.May.28.;.***

Parcy F. 2005. Flowering: a time for integration. ***Int.J.Dev.Biol.2005.;49.(5-6.):585.-93.*** **49: 585-593.**

Parcy F, Nilsson O, Busch MA, Lee I, Weigel D. 1998. A genetic framework for floral patterning. ***Nature*** **395: 561-566.**

Parcy F, Valon C, Kohara A, Misera S, Giraudat J. 1997. The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of Arabidopsis seed development. ***Plant Cell.*** **9: 1265-1277.**

Park S, Harada JJ. 2008. Arabidopsis embryogenesis. ***Methods Mol.Biol.2008.;427.:3-16.*** **427: 3-16.**

Penfield S, Li Y, Gilday AD, Graham S, Graham IA. 2006. Arabidopsis ABA INSENSITIVE4 regulates lipid mobilization in the embryo and reveals repression of seed germination by the endosperm. ***Plant Cell.2006.Aug.;18.(8.):1887.-99.Epub.2006.Jul.14.*** **18: 1887-1899.**

Peng Y, Lin W, Cai W, Arora R. 2007. Overexpression of a Panax ginseng tonoplast aquaporin alters salt tolerance, drought tolerance and cold acclimation ability in transgenic Arabidopsis plants. ***Planta.2007.Aug.;226.(3):729.-40.Epub.2007.Apr*** **226: 729-740.**

Pietersen AM, van Lohuizen M. 2008. Stem cell regulation by polycomb repressors: postponing commitment. ***Curr.Opin.Cell Biol.2008.Apr;20.(2):201.-7.Epub.2008.Mar.4.*** **20: 201-207.**

Poethig RS, Peragine A, Yoshikawa M, Hunter C, Willmann M, Wu G. 2006. The function of RNAi in plant development. ***Cold Spring Harb.Symp.Quant.Biol.2006.;71.:165.-70.*** **71: 165-170.**

Puangsomlee P. 1997. ***Characterisation, cloning and expression studies of CURLY LEAF, a gene involved in leaf and flower development.,*** **University of East Anglia.**

Putterill J, Laurie R, Macknight R. 2004. It's time to flower: the genetic control of flowering time. ***Bioessays.2004.Apr;26.(4):363.-73.*** **26: 363-373.**

Qian C, Zhou MM. 2006. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol.Life Sci.2006.Dec.;63.(23):2755.-63.* 63: 2755-2763.

Qian S, Capovilla M, Pirrotta V. 1993. Molecular mechanisms of pattern formation by the BRE enhancer of the Ubx gene. *EMBO J.* 12: 3865-3877.

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature.2000.Aug.10;406.(6796.):593.-9.* 406: 593-599.

Reinitz J, Levine M. 1990. Control of the initiation of homeotic gene expression by the gap genes giant and tailless in Drosophila. *Dev.Biol.* 140: 57-72.

Ringrose L, Paro R. 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu.Rev.Genet.* 38: 413-443.

Ringrose L, Paro R. 2007. Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* 134: 223-232.

Ringrose L, Rehmsmeier M, Dura JM, Paro R. 2003. Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. *Dev.Cell* 5: 759-771.

Robles P, Pelaz S. 2005. Flower and fruit development in Arabidopsis thaliana. *Int.J.Dev.Biol.2005.;49.(5-6.):633.-43.* 49: 633-643.

Rook F, Hadingham SA, Li Y, Bevan MW. 2006. Sugar and ABA response pathways and the control of gene expression. *Plant Cell Environ.2006.Mar.;29.(3):426.-34.* 29: 426-434.

Ruiz-Garcia L, Madueno F, Wilkinson M, Haughn G, Salinas J, Martinez-Zapater JM. 1997. Different roles of flowering-time genes in the activation of floral initiation genes in Arabidopsis. *Plant Cell.* 9: 1921-1934.

Sabatini S, Heidstra R, Wildwater M, Scheres B. 2003. SCARECROW is involved in positioning the stem cell niche in the Arabidopsis root meristem. *Genes Dev.2003.Feb.1;17(3):354.-8.* 17: 354-358.

Sablowski R. 2007a. Flowering and determinacy in Arabidopsis. *J.Exp.Bot.2007.;58.(5):899.-907.Epub.2007.Feb.10.* 58: 899-907.

Sablowski R. 2007b. The dynamic plant stem cell niches. *Curr.Opin.Plant Biol.2007.Dec.;10(6.):639.-44.Epub.2007.Aug.9.* 10: 639-644.

Salchert K, Bhalerao R, Koncz-Kalman Z, Koncz C. 1998. Control of cell elongation and stress responses by steroid hormones and carbon catabolic repression in plants. *Philos.Trans.R.Soc.Lond B Biol.Sci.* 353: 1517-1520.

Santos MM, Dubreucq B, Miquel M, Caboche M, Lepiniec L. 2005. LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS Lett.2005.Aug.29.;579.(21):4666.-70.* 579: 4666-4670.

Sarkar AK, Luijten M, Miyashima S, Lenhard M, Hashimoto T, Nakajima K, Scheres B, Heidstra R, Laux T. 2007. Conserved factors regulate signalling in Arabidopsis thaliana shoot and root stem cell organizers. *Nature.2007.Apr 12.;446.(7137.):811.-4.* 446: 811-814.

Schiefelbein JW, Somerville C. 1990. Genetic Control of Root Hair Development in Arabidopsis thaliana. *Plant Cell.* 2: 235-243.

Schmitt S, Prestel M, Paro R. 2005. Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev.* 19: 697-708.

Schonrock N, Bouveret R, Leroy O, Borghi L, Kohler C, Gruissem W, Hennig L. 2006. Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway. *Genes Dev.2006.Jun.15.;20.(12.):1667.-78.* 20: 1667-1678.

Schubert D, Clarenz O, Goodrich J. 2005. Epigenetic control of plant development by Polycomb-group proteins. *Curr.Opin.Plant Biol.* 8: 553-561.

Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, Goodrich J. 2006. Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J.* 25: 4638-4649.

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. 2007. Genome regulation by polycomb and trithorax proteins. *Cell* 128: 735-745.

Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V. 2006. Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat.Genet.2006.Jun.;38.(6.):700.-5.Epub.2006.May.28.* 38: 700-705.

Schwartz YB, Pirrotta V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. *Nat.Rev.Genet.2007.Jan.;8.(1):9.-22.* 8: 9-22.

Schwarz S, Grande AV, Bujdoso N, Saedler H, Huijser P. 2008. The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. *Plant Mol.Biol.2008.Feb.17;.* 48: 511-521.

Scott MP, O'Farrell PH. 1986. Spatial programming of gene expression in early Drosophila embryogenesis. ***Annu.Rev.Cell Biol.*** **2: 49-80.**

Shani E, Yanai O, Ori N. 2006. The role of hormones in shoot apical meristem function. ***Curr.Opin.Plant Biol.2006.Oct.;9.(5):484.-9.Epub.2006.Jul.28.*** **9: 484-489.**

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES. 1999. The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. ***Plant Cell*** **11: 445-458.**

Sheldon CC, Finnegan EJ, Rouse DT, Tadege M, Bagnall DJ, Helliwell CA, Peacock WJ, Dennis ES. 2000. The control of flowering by vernalization. ***Curr.Opin.Plant Biol.2000.Oct.;3(5):418.-22.*** **3: 418-422.**

Shimell MJ, Simon J, Bender W, O'Connor MB. 1994. Enhancer point mutation results in a homeotic transformation in Drosophila. ***Science.*** **264: 968-971.**

Simon J. 1995. Locking in stable states of gene expression: transcriptional control during Drosophila development. ***Curr.Opin.Cell Biol.*** **7: 376-385.**

Simon JA, Tamkun JW. 2002. Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. ***Curr.Opin.Genet.Dev.2002.Apr;12.(2):210.-8.*** **12: 210-218.**

Simpson GG. 2004. The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. ***Curr.Opin.Plant Biol.2004.Oct.;7.(5):570.-4.*** **7: 570-574.**

Simpson GG, Gendall AR, Dean C. 1999. When to switch to flowering. ***Annu.Rev.Cell Dev.Biol.*** **15: 519-550.**

Sipos L, Kozma G, Molnar E, Bender W. 2007. In situ dissection of a Polycomb response element in Drosophila melanogaster. ***Proc.Natl.Acad.Sci.U.S.A.2007.Jul.24.;104.(30.):12416.-21.Epub.2007.Jul.18.*** **104: 12416-12421.**

Soderman EM, Brocard IM, Lynch TJ, Finkelstein RR. 2000. Regulation and function of the Arabidopsis ABA-insensitive4 gene in seed and abscisic acid response signaling networks. ***Plant Physiol.2000.Dec.;124.(4):1752.-65.*** **124: 1752-1765.**

Spillane C, MacDougall C, Stock C, Kohler C, Vielle-Calzada JP, Nunes SM, Grossniklaus U, Goodrich J. 2000. Interaction of the Arabidopsis polycomb group proteins FIE and MEA mediates their common phenotypes. ***Curr.Biol.*** **10: 1535-1538.**

Spillane C, Schmid KJ, Laoueille-Duprat S, Pien S, Escobar-Restrepo JM, Baroux C, Gagliardini V, Page DR, Wolfe KH, Grossniklaus U. 2007. Positive darwinian selection at the imprinted MEDEA locus in plants. ***Nature.2007.Jul.*** **448: 349-352.**

Stokes TL, Kunkel BN, Richards EJ. 2002. Epigenetic variation in Arabidopsis disease resistance. ***Genes Dev.2002.Jan.15.;16.(2):171.-82.*** **16: 171-182.**

Stokes TL, Richards EJ. 2002. Induced instability of two Arabidopsis constitutive pathogen-response alleles. ***Proc.Natl.Acad.Sci.U.S.A.2002.May.28.;99.(11):7792.-6.*** **99: 7792-7796.**

Stone SL, Braybrook SA, Paula SL, Kwong LW, Meuser J, Pelletier J, Hsieh TF, Fischer RL, Goldberg RB, Harada JJ. 2008. Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. ***Proc.Natl.Acad.Sci.U.S.A.2008.Feb.26.;105.(8.):3151.-6.Epub.2008.Feb.*** **105: 3151-3156.**

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB, Harada JJ. 2001. LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. ***Proc.Natl.Acad.Sci.U.S.A.2001.Sep.25.;98.(20.):11806.-11.*** **98: 11806-11811.**

Sun FL, Cuaycong MH, Elgin SC. 2001. Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. ***Mol.Cell Biol.2001.Apr;21(8.):2867.-79.*** **21: 2867-2879.**

Sun J, Niu QW, Tarkowski P, Zheng B, Tarkowska D, Sandberg G, Chua NH, Zuo J. 2003. The Arabidopsis AtIPT8/PGA22 gene encodes an isopentenyl transferase that is involved in de novo cytokinin biosynthesis. ***Plant Physiol.2003.Jan.;131.(1):167.-76.*** **131: 167-176.**

Sung S, Amasino RM. 2004. Vernalization and epigenetics: how plants remember winter. ***Curr.Opin.Plant Biol.*** **7: 4-10.**

Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, Goto K, Jacobsen SE, Amasino RM. 2006. Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. ***Nat.Genet.*** **38: 706-710.**

Suzuki M, Wang HH, McCarty DR. 2007. Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. ***Plant Physiol.2007.Feb.;143.(2):902.-11.Epub.2006.Dec.8.*** **143: 902-911.**

Swarup R, Parry G, Graham N, Allen T, Bennett M. 2002. Auxin cross-talk: integration of signalling pathways to control plant development. ***Plant Mol.Biol.2002.Jun.-Jul.;49.(3-4):411.-26.*** **49: 411-426.**

Tadege M, Sheldon CC, Helliwell CA, Upadhyaya NM, Dennis ES, Peacock WJ. 2003. Reciprocal control of flowering time by OsSOC1 in transgenic Arabidopsis and by FLC in transgenic rice. ***Plant Biotechnol.J.2003.Sep.;1(5):361.-9.*** **1: 361-369.**

Tanaka M, Kikuchi A, Kamada H. 2008. The Arabidopsis Histone Deacetylases HDA6 and HDA19 Contribute to the Repression of Embryonic Properties after Germination. ***PLANT PHYSIOLOGY*** **146: 149-161.**

Tang X, Hou A, Babu M, Nguyen V, Hurtado L, Lu Q, Reyes JC, Wang A, Keller WA, Harada JJ, Tsang EW, Cui Y. 2008. The Arabidopsis BRAHMA Chromatin Remodelling ATPase Is Involved in Repression of Seed Maturation Genes in Leaves. ***Plant Physiol.2008.May.28.;.***

Tariq M, **Saze H**, **Probst AV**, **Lichota J**, **Habu Y**, **Paszkowski J**. 2003**Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin.** Proc Natl Acad Sci U S A. Jul 22;100(15):8823-7. Epub 2003 Jul 9.

Teale WD, Paponov IA, Palme K. 2006. Auxin in action: signalling, transport and the control of plant growth and development. ***Nat.Rev.Mol.Cell Biol.2006.Nov.;7.(11):847.-59.Epub.2006.Sep.20.*** **7: 847-859.**

Telfer A, Bollman KM, Poethig RS. 1997. Phase change and the regulation of trichome distribution in Arabidopsis thaliana. ***Development.*** **124: 645-654.**

Telfer A, Poethig RS. 1998. HASTY: a gene that regulates the timing of shoot maturation in Arabidopsis thaliana. ***Development.*** **125: 1889-1898.**

Tie F, Furuyama T, Prasad-Sinha J, Jane E, Harte PJ. 2001. The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. ***Development*** **128: 275-286.**

Tohno Y, Tohno S. 1993. A possible association of nuclear lamina with chromatin. ***Cell Mol.Biol.(Noisy.-le-grand).*** **39: 757-764.**

Tolhuis B, Muijrers I, de Wit E, Teunissen H, Talhout W, van Steensel B, van Lohuizen M. 2006. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. ***Nat.Genet.*** **38: 694-699.**

Tsukagoshi H, Morikami A, Nakamura K. 2007. Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in

Arabidopsis seedlings. *Proc.Natl.Acad.Sci.U.S.A.2007.Feb.13;104.(7.):2543.-7.Epub.2007.Jan.31.* 104: 2543-2547.

Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V. 2007. Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS.Genet.2007.Jun.;3(6.):e86.Epub.2007.Apr 17.* 3: e86.

Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ. 1997. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell.* 9: 1963-1971.

van den Berg C, Willemsen V, Hage W, Weisbeek P, Scheres B. 1995. Cell fate in the Arabidopsis root meristem determined by directional signalling. *Nature* 378: 62-65.

van den Berg C, Willemsen V, Hendriks G, Weisbeek P, Scheres B. 1997. Short-range control of cell differentiation in the Arabidopsis root meristem. *Nature* 390: 287-289.

Wagner D, Sablowski RW, Meyerowitz EM. 1999. Transcriptional activation of APETALA1 by LEAFY. *Science.* 285: 582-584.

Wang D, Tyson MD, Jackson SS, Yadegari R. 2006. Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis.
***Proc.Natl.Acad.Sci.U.S.A.2006.Aug.29.;103.(35.):13244.-9.Epub.2006.Aug.21.* 103: 13244-13249.**

Wang JW, Schwab R, Czech B, Mica E, Weigel D. 2008. Dual Effects of miR156-Targeted SPL Genes and CYP78A5/KLUH on Plastochron Length and Organ Size in Arabidopsis thaliana. *Plant Cell.2008.May.20.;.*

Waterhouse PM, Wang MB, Lough T. 2001. Gene silencing as an adaptive defence against viruses. *Nature.2001.Jun.14;411.(6839.):834.-42.* 411: 834-842.

Weigel D, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM. 1992. LEAFY controls floral meristem identity in Arabidopsis. *Cell* 69: 843-859.

Weigel D, Nilsson O. 1995. A developmental switch sufficient for flower initiation in diverse plants. *Nature.* 377: 495-500.

Welch D, Hassan H, Blilou I, Immink R, Heidstra R, Scheres B. 2007. Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action. *Genes Dev.2007.Sep.1;21(17):2196.-204.* 21: 2196-2204.

West M, Yee KM, Danao J, Zimmerman JL, Fischer RL, Goldberg RB, Harada JJ. 1994. LEAFY COTYLEDON1 Is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in Arabidopsis. ***Plant Cell.*** **6: 1731-1745.**

White RA, Lehmann R. 1986. A gap gene, hunchback, regulates the spatial expression of Ultrabithorax. ***Cell.*** **47: 311-321.**

Willemsen V, Scheres B. 2004. Mechanisms of pattern formation in plant embryogenesis. ***Annu.Rev.Genet.2004.;38.:587.-614.*** **38: 587-614.**

William DA, Su Y, Smith MR, Lu M, Baldwin DA, Wagner D. 2004. Genomic identification of direct target genes of LEAFY. ***Proc.Natl.Acad.Sci.U.S.A.2004.Feb.10;101.(6.):1775.-80.Epub.2004.Jan.21.*** **101: 1775-1780.**

Wobus U, Weber H. 1999. Seed maturation: genetic programmes and control signals. ***Curr.Opin.Plant Biol.*** **2: 33-38.**

Woo YM, Park HJ, Su'udi M, Yang JI, Park JJ, Back K, Park YM, An G. 2007. Constitutively wilted 1, a member of the rice YUCCA gene family, is required for maintaining water homeostasis and an appropriate root to shoot ratio. ***Plant Mol.Biol.2007.Sep.;65.(1-2):125.-36.Epub.2007.Jul.6.*** **65: 125-136.**

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA. 2006. The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. ***Proc.Natl.Acad.Sci.U.S.A*** **103: 14631-14636.**

Wu G, Poethig RS. 2006. Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. ***Development.2006.Sep.;133.(18.):3539.-47.Epub.2006.Aug.16.*** **133: 3539-3547.**

Yadegari R, Kinoshita T, Lotan O, Cohen G, Katz A, Choi Y, Katz A, Nakashima K, Harada JJ, Goldberg RB, Fischer RL, Ohad N. 2000. Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. ***Plant Cell*** **12: 2367-2382.**

Yamamoto K, Sonoda M, Inokuchi J, Shirasawa S, Sasazuki T. 2004. Polycomb Group Suppressor of Zeste 12 Links Heterochromatin Protein 1{alpha} and Enhancer of Zeste 2. ***J.Biol.Chem.*** **279: 401-406.**

Yang M, Sack FD. 1995. The too many mouths and four lips mutations affect stomatal production in Arabidopsis. ***Plant Cell.*** **7: 2227-2239.**

Yi H, Richards EJ. 2008. Phenotypic instability of Arabidopsis alleles affecting a disease Resistance gene cluster. ***BMC.Plant Biol.2008.Apr 14;8.:36.*** **8: 36.**

Yu J, Yu J, Rhodes DR, Tomlins SA, Cao X, Chen G, Mehra R, Wang X, Ghosh D, Shah RB, Varambally S, Pienta KJ, Chinnaiyan AM. 2007. A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. *Cancer Res.2007.Nov.15.;67.(22.):10657.-63.* 67: 10657-10663.

Zhang CC, Bienz M. 1992. Segmental determination in Drosophila conferred by hunchback (hb), a repressor of the homeotic gene Ultrabithorax (Ubx). *Proc.Natl.Acad.Sci.U.S.A.* 89: 7511-7515.

Zhang H, Rider SD, Jr., Henderson JT, Fountain M, Chuang K, Kandachar V, Simons A, Edenberg HJ, Romero-Severson J, Muir WM, Ogas J. 2008. The CHD3 remodeler PICKLE promotes trimethylation of histone H3 lysine 27. *J.Biol.Chem.2008.Jun.6.;*

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007. Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis. *PLoS.Biol.* 5: e129.

Zhou F, Mosher S, Tian M, Sassi G, Parker J, Klessig DF. 2008. The Arabidopsis gain-of-function mutant ssi4 requires RAR1 and SGT1b differentially for defense activation and morphological alterations. *Mol.Plant Microbe Interact.2008.Jan.;21(1):40.-9.* 21: 40-49.

Zik M, Irish VF. 2003. Flower development: initiation, differentiation, and diversification. *Annu.Rev.Cell Dev.Biol.2003.;19.:119.-40.* 19: 119-140.

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W. 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* 136: 2621-2632.

Zink B, Engstrom Y, Gehring WJ, Paro R. 1991. Direct interaction of the Polycomb protein with Antennapedia regulatory sequences in polytene chromosomes of Drosophila melanogaster. *EMBO J.* 10: 153-162.

8. 0. Appendix

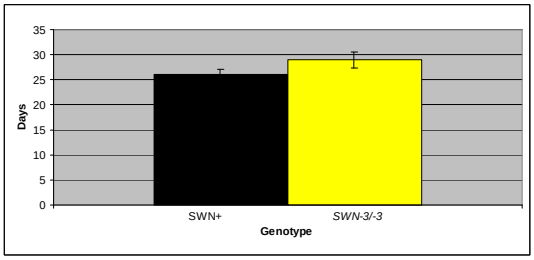

Fig. 8. 1. The effect of *swn-3/-3* on flowering time in long days at 30 °C.
The flowering time (number of days for a 1 m bolt to form) of *swn-3/-3* and *SWN+* plants in long days at 30 °C. There was no difference in number of days it took to flower between *swn-3/-3* and *SWN+*. (n = 30).

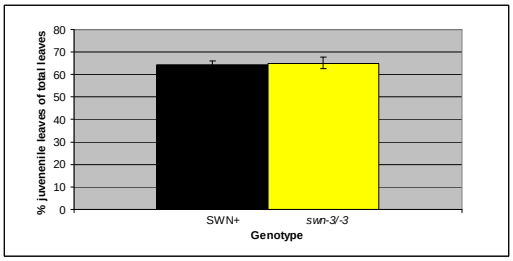

Fig. 8. 2. The effect of *swn-3/-3* on percentage of juvenile leaves in long days.
The average of percentage of juvenile leaves of 11 families of *swn-3/-3* and *SWN+* derived from the same F2 population showed no significant difference. (n = 11 families, 12 plants family^{-1}).

The most up regulated genes in *swn-7*		
Fold Change	**Gene Name**	**Gene Title**
14.3	At5g02540	short-chain dehydrogenase/reductase (SDR) family protein
8.3	At2g17880	DNAJ heat shock protein, putative
7.7	At5g04190	phytochrome kinase substrate-related
7.7	At1g56150	auxin-responsive family protein
6.7	At3g62950	glutaredoxin family protein
6.7	At1g06080	delta 9 desaturase (ADS1)
6.25	At5g50335	expressed protein
6.25	At3g12900	oxidoreductase, 2OG-Fe(II) oxygenase family protein
5.5	At5g53980	homeobox-leucine zipper family protein
5.5	At4g32280	auxin-responsive AUX/IAA family protein
5.5	At2g19970	pathogenesis-related protein, putative
5.5	At5g15160	bHLH family protein
5.3	At3g53250	auxin-responsive family protein
5	At5g51810	gibberellin 20-oxidase, putative
5	At1g20190	expansin, putative (EXP11)
5	At1g28330	dormancy-associated protein, putative (DRM1)
4.8	At5g07010	sulfotransferase family protein
4.8	At5g04150	basic helix-loop-helix (bHLH) family protein
4.5	At3g60530	zinc finger (GATA type) family protein
4.5	At5g61160	transferase family protein
4.5	At1g49860	glutathione S-transferase, putative
4.3	At1g69490	no apical meristem (NAM) family protein
4.3	At4g29190	zinc finger (CCCH-type) family protein
4.2	At1g79700	ovule development protein, putative
4	At3g15450	expressed protein
4	At2g37640	expansin, putative (EXP3)
4	At2g14580	pathogenesis-related protein, putative
4	At4g35770	senescence-associated protein (SEN1)
4	At2g41230	expressed protein
3.8	At1g49210	zinc finger (C3HC4-type RING finger) family protein
The most down regulated genes in *swn-7*		
Fold Change	**Gene Name**	**Gene Title**
24.76	At5g44120	12S seed storage protein (CRA1)
23.62	At4g27140	2S seed storage protein 1
22.05	At4g28520	12S seed storage protein, putative / cruciferin, putative
21.32	At5g48850	male sterility MS5 family protein
15.92	At4g27160	2S seed storage protein 3
14.93	At4g27150	2S seed storage protein
11.78	At1g03880	12S seed storage protein (CRB)
11.7	At4g25140	glycine-rich protein / oleosin
10.74	At2g44460	glycosyl hydrolase family 1 protein
10.6	At2g27690	cytochrome P450, putative

10.3	At5g24660	expressed protein
9.88	At2g14560	expressed protein
9.876	At3g49580	expressed protein
9.508	At5g59310	lipid transfer protein 4 (LTP4)
9.069	At3g15310	expressed protein
8.802	At4g31800	WRKY family transcription factor
8.761	At2g34600	expressed protein
8.736	At2g39330	jacalin lectin family protein
8.171	At5g66400	dehydrin (RAB18)
8.053	At5g03350	legume lectin family protein
7.651	At1g72520	lipoxygenase, putative
7.178	At5g40420	glycine-rich protein / oleosin
7.157	At5g26220	ChaC-like family protein
7.115	At1g17380	expressed protein
6.516	At3g22740	homocysteine S-methyltransferase 3 (HMT-3)
6.452	At3g47960	proton-dependent oligopeptide transport (POT) family protein
6.442	At4g22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
6.222	At1g17420	lipoxygenase, putative
6.215	At4g21680	proton-dependent oligopeptide transport (POT) family protein

The most up regulated genes in *clf-28*		
Fold Change	**Gene Name**	**Gene Title**
50	At4g18960	floral homeotic protein AGAMOUS (AG)
25	At5g50790	nodulin MtN3 family protein
16.7	At1g09500	cinnamyl-alcohol dehydrogenase family / CAD family
16.7	At1g24260	MADS-box protein (AGL9)
16.7	At1g69490	no apical meristem (NAM) family protein
14.3	At2g39510	nodulin MtN21 family protein
14.3	At3g19550	expressed protein
12.5	At4g36740	homeobox-leucine zipper family protein
12.5	At5g39520	expressed protein
12.5	At1g80160	lactoylglutathione lyase family protein / glyoxalase I family protein
12.5	At5g24910	cytochrome P450 family protein
12.5	At1g03710	expressed protein
10	At2g18550	homeobox-leucine zipper family protein
10	At5g39220	hydrolase, alpha/beta fold family protein
9.6	At1g65480	flowering locus T protein (FT)
9.6	At5g51810	gibberellin 20-oxidase, putative
8.9	At5g60140	transcriptional factor B3 family protein
8.3	At3g46660	UDP-glucoronosyl/UDP-glucosyl transferase family protein
8.3	At2g42830	agamous-like MADS box protein AGL5 / floral homeodomain transcription factor (AGL5)
8.3	At1g16950	expressed protein
7.5	At3g20210	vacuolar processing enzyme, putative / asparaginyl endopeptidase, putative
7.5	At4g36700	cupin family protein

7.5	At5g10140	MADS-box protein flowering locus F (FLC)
7.5	At5g04150	basic helix-loop-helix (bHLH) family protein
7.5	At5g02020	expressed protein
6	At5g57530	endo-xyloglucosyl transferase, putative / xyloglucan endotransglycosylase
5.5	At2g20700	expressed protein
5.25	At5g22430	expressed protein
5	At2g17880	DNAJ heat shock protein, putative
5	At4g35690	hypothetical protein
The most down regulated genes in *clf-28*		
Fold Change	**Gene Name**	**Gene Title**
33.44	AT4G28520	12S seed storage protein, putative / cruciferin, putative
30.35	AT5G44120	12S seed storage protein (CRA1)
20.04	AT4G27140	2S seed storage protein 1 / 2S albumin storage protein / NWMU1-2S albumin 1
18.27	AT5G48850	male sterility MS5 family protein
15.99	AT4G27160	2S seed storage protein 3 / 2S albumin storage protein / NWMU2-2S albumin 3
15.43	AT5G03350	legume lectin family protein
13.23	AT4G27150	2S seed storage protein 2 / 2S albumin storage protein / NWMU2-2S albumin 2
12.76	AT1G03880	12S seed storage protein (CRB)
11.22	AT2G27690	cytochrome P450, putative
10.69	AT2G14560	expressed protein
10.04	AT5G40420	glycine-rich protein / oleosin
9.543	AT4G14400	ankyrin repeat family protein
9.482	AT3G22231	expressed protein
8.665	AT2G34600	expressed protein
7.784	AT3G15310	expressed protein
7.505	AT4G31800	WRKY family transcription factor
7.442	AT4G25140	glycine-rich protein / oleosin
7.426	AT2G24600	ankyrin repeat family protein
7.414	AT4G27170	2S seed storage protein 4 / 2S albumin storage protein / NWMU2-2S albumin 4
7.361	AT1G17420	lipoxygenase, putative
7.012	AT5G51720	expressed protein
6.847	AT4G14365	zinc finger (C3HC4-type RING finger) family protein / ankyrin repeat family protein
6.782	AT1G72520	lipoxygenase, putative
6.409	AT5G13220	expressed protein
6.311	AT5G66400	dehydrin (RAB18)
6.195	AT2G15010	thionin, putative
6.165	AT3G06070	expressed protein
6.147	AT3G49580	expressed protein
6.087	AT1G43590	hypothetical protein
5.936	AT5G26220	ChaC-like family protein

138.3	At3g27690	chlorophyll A-B binding protein (LHCB2:4)
126.1	At2g18300	basic helix-loop-helix (bHLH) family protein
118.1	At1g66100	thionin, putative
116.9	At1g32480	membrane protein, putative
114.3	At1g73610	late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing protein
113.6	At2g29290	tropinone reductase, putative / tropine dehydrogenase, putative
113.1	At3g15670	late embryogenesis abundant protein, putative / LEA protein, putative
110.2	At5g48490	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
97.1	At5g62490	ABA-responsive protein (HVA22b)
97.09	At5g44020	acid phosphatase class B family protein
93.57	At5g54940	cysteine proteinase, putative
91.89	At4g26530	fructose-bisphosphate aldolase, putative
90.5	At1g54870	short-chain dehydrogenase/reductase (SDR) family protein
89.57	At3g01500	carbonic anhydrase 1, chloroplast / carbonate dehydratase 1 (CA1)
88.6	At1g75830	plant defensin-fusion protein, putative (PDF1.1)
88.04	At3g06070	expressed protein
86.36	At2g25890	glycine-rich protein / oleosin
82.71	At1g20440	auxin-responsive family protein
79.98	At5g07330	expressed protein
79.7	At3g46780	expressed protein
79.7	At2g41260	glycine-rich protein / late embryogenesis abundant protein (M17)
1020.4	At5g14740	carbonic anhydrase 2 / carbonate dehydratase 2 (CA2) (CA18)
970.9	At2g10940	late embryogenesis abundant group 1 domain-containing protein / LEA group 1 domain-containing protein
961.5	At1g32560	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
900.9	At4g22490	mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23
900.9	At1g21500	expressed protein
885	At4g16160	expressed protein
869.6	At2g42530	cold-responsive protein / cold-regulated protein (cor15b)
847.5	At1g05510	expressed protein
833.3	At5g16030	expressed protein
793.7	At3g51810	EMP-like protein GEA1 (EM1)
775.2	At1g47980	expressed protein
763.4	At1g48130	peroxiredoxin (PER1) / rehydrin, putative
740.7	At5g66780	expressed protein
719.4	At2g21490	dehydrin family protein
709.2	At1g04560	AWPM-19-like membrane family protein
689.7	At4g26740	embryo-specific protein 1 (ATS1)
680.3	At2g28490	cupin family protein

The most up regulated genes in *swn-7 clf-28*

Fold Change	Gene Name	Gene Title
2961.1	At5g45690	expressed protein
2960.7	At3g21380	jacalin lectin family protein
1785.7	At2g36640	late embryogenesis abundant protein (ECP63) / LEA protein
1724.1	At1g47540	trypsin inhibitor, putative
1724.1	At3g21370	glycosyl hydrolase family 1 protein
1612.9	At3g17520	late embryogenesis abundant domain-containing protein / LEA domain-containing protein
1515.2	At3g53040	late embryogenesis abundant protein, putative / LEA protein, putative
1470.6	At4g27170	2S seed storage protein 4 / 2S albumin storage protein / NWMU2-2S albumin 4
1449.3	At4g09600	gibberellin-regulated protein 3 (GASA3) / gibberellin-responsive protein 3
1449.3	At2g27380	proline-rich family protein
1331.3	At1g32560	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
1087	At4g16160	expressed protein
1087	At1g05510	expressed protein
1041.7	At3g51810	EMP-like protein GEA1 (EM1)

The most down regulated genes in *swn-7 clf-28*

Fold Change	Gene Name	Gene Title
241.8	At3g22231	expressed protein
167.7	At3g05730	expressed protein
162.8	At1g31580	expressed protein
161.2	At1g29430	auxin-responsive family protein
156.1	At1g29660	GDSL-motif lipase/hydrolase family protein
145.8	At3g14210	myrosinase-associated protein, putative

Table. 8. 1. Sample of mis-expressed genes in the Pc-G mutants.
The 30 most highly mis-expressed genes in *swn-7*, *clf-28*, and *swn-7 clf-28*. The level of mis-expression is the fold change is relative to Col (wild-type) signal intensity. Gene titles are derived from GO terms from the *Arabidopsis* Information Resources which were assigned by the Affymetric service.

Table. 8. 2. Discrete, redundant, and common *SWN* and CLF targets, both direct and indirect.
See attached CD (file Table. 8. 2.) Gene found up regulated in *swn-7*, *clf-28*, and *swn-7 clf-28* where compared to genes known to possess H3K27me^3. Genes up regulated that possess H3K27me^3 are considered as direct targets (indicated in bold), genes not possessing H3K27me^3 are considered as indirect targets. Genes up regulated in both *swn-7* and *swn-7 clf-28* are considered as discrete *SWN* targets (found on the tab "Discrete *SWN* targets). Genes up regulated in both *clf-28* and *swn-7 clf-28* are considered as discrete *CLF* targets (found on the tab "Discrete *CLF* targets). Genes up regulated in *swn-7* and *clf-28* were considered as common targets for *SWN* and *CLF* (found on the tab "common *SWN CLF* targets). Genes found only up regulated in *swn-7 clf-28* mutants were considered as redundant *SWN CLF* targets (found on the tab "Redundant *SWN CLF* targets). microarray analysis can be found in "Materials and Methods" section. The direct, and indirect, target gene lists were sorted by AGI number, and gene titles were designated by the www.virtualplant.org, based on the *Arabidopsis* Information Resource Gene Ontology terms.

Table. 8. 3. Common direct targets of the Pc-G the members.
See attached CD (file Table. 8. 3.). Genes found commonly up regulated in either *clf-28*, or, *swn-7 clf-28*, and *emf2-1*, or *msi1*- were analyzed to find which possessed H3K27me^3. Genes found commonly up regulated in two mutants and possessed H3K27me^3 were considered as common direct target, i.e. direct Pc-G targets. For example, common targets of CLF and EMF2 are in the tab "CLF EMF2 direct targets" and SWN CLF and EMF2 targets are in the tab labelled "SWN CLF and EMF2 direct targets" etc…The gene lists were sorted by AGI number, and gene titles were designated by the www.virtualplant.org, based on the *Arabidopsis* Information Resource Gene Ontology terms. Gene lists analysis was carried out using "Sungear" function on the www.virtualplant.org.

Table. 8. 4. Pc-G targets up regulated in *hsi1- hsl2*, or *pkl-*, or *emf1-* mutants.
See attached CD (file Table. 8. 4.). Genes found commonly up regulated in either *clf-28*, or, *swn-7 clf-28*, and *hsi1- hsl2-*, or *pkl-1*, or *emf1-* were analyzed to find which possessed H3K27me^3. Genes found commonly up regulated in two mutants and possessed H3K27me^3 were considered as common direct target, i.e. direct Pc-G targets. For example, common targets of *swn-7 clf-28* and *PKL* are in the tab "Pc-G and *PKL* targets", and CLF and EMF1 targets are in the tab labelled "CLF and EMF1 targets" etc… The gene lists were sorted by AGI number, and gene titles were designated by the www.virtualplant.org, based on the *Arabidopsis* Information Resource Gene Ontology terms. Gene lists analysis was carried out using "Sungear" function on the www.virtualplant.org.

Table. 8. 5. Common direct targets of the Pc-G and TFL2.
See attached CD (file Table. 8. 5.). Gene found up regulated in *swn-7 clf-28* were compared to genes found by TFL2 which possessed H3K27me^3. These genes are considered as common targets of the Pc-G and *TFL2*. The gene lists were sorted by AGI number, and gene titles were designated by the www.virtualplant.org, based on the *Arabidopsis* Information Resource Gene Ontology terms. Gene lists analysis was carried out using "Sungear" function on the www.virtualplant.org.